

“The *Government* Beguiled Me”:
The Entrapment Defense and the Problem of Private Entrapment

By Gideon Yaffe

Introduction

In the United States, a criminal defendant can show himself to be not guilty of the crime of which he is accused by showing that he was “entrapped” by agents of the government. Entrapment is not merely a mitigating factor to be considered in sentencing a defendant who is acknowledged to be guilty of a crime; it is not treated in the law in the way, for instance, a defendant’s poverty is often treated, as providing reason to give a guilty defendant a lighter sentence than an equally guilty but rich defendant, who also stole for money. Rather, entrapment constitutes a complete defense: An entrapped defendant is simply not guilty of violating the law under which he is being prosecuted. Those who think that there would be something wrong with a legal system that did not allow the entrapment defense have, in the background of their thoughts, a particular picture of the conditions under which it is appropriate to hold a person legally responsible for his behavior, conditions that are absent when that person has given in to certain temptations supplied by the government. They accept, that is, however implicitly, a theory of legal responsibility under which certain people who, under certain circumstances, give in to a temptation issued by the government are thereby, and on those grounds alone, rightly excused from legal responsibility. This paper describes some of the features of such a theory, a theory from which it follows that the entrapment defense is an essential element in a just legal system.

In the very first case in which a U.S. court considered the entrapment defense, *Board of Commissioners v. Backus* (29 How. Pr. 33, 42 (1864)), a New York court rejected it on the grounds that God refused to excuse Eve when she opined, “The serpent beguiled me, and I did eat.”¹ If the excuse of “beguilement” wasn’t good enough for God, why should it be good enough for the court?² In offering this justification for its action, the court chose to interpret the defense as claiming, generally, that “beguiled” defendants are not guilty, a claim it took to be false. In fact, however, defendants who offer the defense are claiming not just to have been beguiled, but also to have been beguiled *by the government*. The court assumed that this aspect of the defense was irrelevant. They assumed, that is, that government beguilement is no

more an excuse than serpent. Call this "the problem of private entrapment": A justification of the entrapment defense must explain why it is that, had a defendant who successfully employs the entrapment defense been tempted by a private party, instead of the government, he would not have been excused from legal responsibility.³ In this first entrapment case, then, the court assumed that there could be no solution to the problem of private entrapment, and so rejected the defendant's argument.⁴

With the goal of describing part of the theory of legal responsibility underlying the entrapment defense, this paper aims to solve the problem of private entrapment.⁵ Imagine a pair of defendants who act precisely the same way in response to a temptation offered by another person whom each believed to be a private party, not working for the government. Perhaps both defendants are approached by someone who claims to be able to provide security codes for a facility holding something worth stealing to someone willing to steal it, and imagine that both defendants, neither of whom has done, or contemplated doing, anything like this before, give in to the temptation and are arrested on leaving the facility with the stolen goods in hand. It will be argued that they differ in responsibility if one of them was tempted by a private party while the other, unbeknownst to him, was tempted by an agent of the government acting with the aim of prosecuting the defendant should he give in to the temptation.⁶

The paper is structured as follows. Section 1 defines various concepts useful for thinking about entrapment and situates the problem of private entrapment within the conceptual framework described. Section 2 draws out the details of the first of two broad strategies for justifying the entrapment defense: an approach according to which entrapment undermines a certain class of defendants' (those who are not "predisposed") culpability for the crime. The section argues that those who adopt this approach are committed to a retributivist justification of legal punishment, in contrast to other sorts. This result opens the door to the possibility that the privately entrapped are, and the governmentally entrapped (and unpredisposed) are not, *deserving* of legal punishment, a claim that will be argued for in section 4. Section 3 turns to the second broad strategy for justifying the defense: the approach according to which the defense is justified because entrapment involves unacceptable action on the part of the government, quite independently of the culpability of the defendant. The *Model Penal Code's* test for entrapment is used as an example of such an approach. The section argues that any adequate effort to pursue this strategy necessarily encounters the problem of private entrapment; the problem is solved from this approach only at the cost of the adequacy of the justification of the defense supplied. Section 4 identifies a necessary condition for desert of legal punishment and argues that the privately entrapped satisfy that condition, as do the predisposed, while the governmentally entrapped and unpredisposed do not. Thus, the problem of private

entrapment is solved. By the end, then, we will have a clearer picture of what one needs to think legal responsibility is, and what sorts of conditions undermine it, if one of the strongest obstacles to the justification of the entrapment defense is to be overcome.

1. *Sharpening the Problem of Private Entrapment*

In all the cases, and hypothetical cases, that concern us, the defendant gave in to a temptation to commit a crime, a temptation provided by another party. This class of cases is not as large as one might think, for there is a difference between taking bait, or merely availing oneself of an opportunity, and giving in to temptation. Say, for instance, that the police leave a door unlocked and lie in wait to see if anyone tries, opportunistically, to steal what is behind it. Someone who takes the bait cannot avail himself of the entrapment defense merely on the grounds that he would not have been able to take the goods had the government not left the door unlocked. In this case, the government does not offer the crime to the defendant *under the description banned by the law*. The government never says to him "Why don't you steal that stuff?" for the purpose of causing him to do so. An opportunity does not count as a temptation to perform a particular type of action, in the sense in which the term "temptation" will be used here, unless it presents the action under the description that defines the relevant type of action. If you tempt someone to steal something then you must present to him the opportunity to steal *as such*. The problem of private entrapment arises, then, for all and only those cases in which we are ready to excuse a defendant for giving in to a temptation, in the sense meant here, when it is supplied by the government, but not when it is supplied by a private party.⁷

The notion of temptation just defined provides a first step, although only a first step, toward distinguishing bogus entrapment defenses from legitimate. A defendant's entrapment defense is to be rejected if the government never represented the act it made available to the defendant under the description banned by the law. However, not every defendant who acts in response to a genuine, government-supplied temptation should have an entrapment defense available. Additional work needs to be done to distinguish among defendants who acted in response to a government-supplied temptation. How you choose to draw the line among these defendants will indicate what you take to be the fundamental basis of the entrapment defense; it will indicate, that is, what you take to be the crucial feature of a case that makes a not-guilty verdict appropriate. Two broad approaches can be taken in drawing this line.

To understand the two approaches for distinguishing among defendants who employ the entrapment defense, and the two associated justifications of it, it helps to have in hand a conception of two parallel aspects of

legal responsibility. Both legally responsible and morally responsible actions are brought about in the familiar way that qualifies them minimally to be held to legal or moral standards, respectively. At the least, this means that the actions for which the agent is being held responsible were not subject to certain pressures, such as coercive pressures, that seem to undermine their capacity to reflect the facts about the agent that we take ourselves to be most concerned with in assessing legal or moral responsibility. It is possible, in fact likely, that we have different concerns in judging someone legally, as opposed to morally responsible, and this difference might manifest itself in a difference in what does and does not count as action that is appropriately held to legal and moral standards, respectively. But still, both legal and moral responsibility assessments are sensitive to the particular way in which the act being assessed came to be performed.

However, in the moral case, the question of whether or not it is appropriate to punish or in some other way censure someone who has acted wrongly does not depend on the behavior of those who are to judge him and administer his punishment. That is, a person has genuinely done wrong and is genuinely morally responsible for his behavior even if he is surrounded solely by such morally unsavory characters that there is nobody pure enough to be in position to heap him with moral outrage, and even if his transgressions have come to light only as a result of morally objectionable spying and squealing. To say that no one is free enough from sin to cast the first stone at the sinner is not to say that he is not a sinner. However, legal responsibility, at least in criminal cases, is different from moral in this respect. We do not hold a person legally responsible if the government has behaved inappropriately toward him in some part of the process that would, given perfect governmental conduct, lead to his punishment. We have laws that tell us when arrest, prosecution, conviction and punishment do and do not proceed legally, and in some cases we do not take defendants to be legally responsible when these procedures are not followed correctly. That is, for the purposes of legal responsibility, and not usually for moral, the conduct of those doing the assessing and the punishing matters. Broadly, then, we can say that a person is legally responsible for his conduct only if (1) he is what I will call "legally culpable" for it — he has acted in a way banned by the law through a process that qualifies him to be held to the law's standards with regard to that conduct⁸; and (2) his treatment by the government with respect to the action has been legally impeccable; it has lived up to the standards of government conduct that the law requires. Substantive criminal defenses are aimed at undermining the prosecution's claim of the defendant's legal culpability; procedural defenses, by contrast, are intended to show that the agent is not legally responsible, even if legally culpable, because the government has not treated him as it ought.

It is difficult to tell whether the entrapment defense is intended to be a substantive or a procedural defense. In acquitting an entrapped defendant, is the court saying that the defendant's agency was undermined in some way such that a statement of the form "He performed an act of type C" is true, but not true in the way that it would need to be to apply the law against acts of that type to his case? Or, rather, is the court saying that some part of the governmental process that would lead to his punishment has been tainted in some objectionable way? Is a court that grants an entrapment defense finding that the defendant did not act badly, by legal standards, or that the government did something it ought not to have done, by legal standards?

In the law, two distinctively different approaches can be found to the question of whether a particular defendant, who performed the acts for which he is being tried as a result of a temptation supplied by the government, was entrapped and is, therefore, not guilty of the crime of which he is accused. There can be found, that is, two different ways of drawing the crucial line between defendants who give in to government-supplied temptation. These two approaches conform precisely to the two aspects just described of legal responsibility under the criminal law. Some have thought that entrapped defendants are simply not legally culpable for what they have done; others have thought, by contrast, that entrapped defendants may be culpable, but are not legally responsible because the government has acted improperly toward them. The former approach is the dominant approach and it has been employed by the majority in each of the six Supreme Court cases concerning entrapment.⁹ By contrast, the minority, in all six cases, advocated the alternative approach. Courts that take entrapment to eliminate culpability ordinarily employ what is sometimes called the "Subjective Test" of entrapment, while courts that take entrapment to excuse because of governmental misconduct employ what is known as the "Objective Test."¹⁰ More about these tests shortly.

Whichever test you advocate, you identify a feature that you claim to be present in all and only those cases of governmentally tempted defendants whose entrapment defenses should be honored — either lack of culpability or governmental misconduct. However, that feature cannot also be present in parallel cases of private temptation, in which the defense is not to be granted, without thereby undermining the justification for allowing the entrapment defense when the government *is* involved. The principle that sorts governmentally tempted defendants — some of their entrapment defenses are to be allowed, some denied — must still allow for lumping of the privately tempted, none of whom should be granted an entrapment defense. This is another way of putting the problem of private entrapment. The test for determining if a governmentally tempted defendant is rightly granted an entrapment defense implies a justification for allowing the defense at all ("such people aren't culpable," or "we don't want our government to act that

way”); this justification will be adequate only if it doesn’t also imply that it would be justified to allow the defense to the privately entrapped.

Notice that there is great initial plausibility to the following two claims. First: *The advocate of the Objective Test gets a solution to the problem of private entrapment for free.* After all, if entrapment defenses that are to be honored all involve some kind of government misconduct, then it is no surprise that the privately tempted are denied the defense; by definition, they are not the victims of *government* misconduct. And, second: *The advocate of the Subjective Test cannot hope to solve the problem of private entrapment.* After all, if unpredisposed defendants are really not culpable when they accept temptations, it shouldn’t matter who is issuing them, and so it seems that they would not be legally responsible if tempted by a private party; this seems especially true given that usually governmentally tempted defendants have no idea that it is the government that is tempting them, but believe themselves, instead, to be tempted by private parties. However, despite their plausibility I will argue that *both of these claims are false.* The argument for this is spread over the remainder of this paper.

2. *The Subjective Test and Retribution*

According to the Subjective Test, the entrapment defense fails if it can be shown that the defendant was “predisposed” to commit the crime. Thus, according to this approach, the fact that the government provided the temptation to commit the crime does not automatically ameliorate the defendant’s culpability, even if the temptation was, by ordinary measures, extremely powerful, and even if the government acted improperly in supplying it; predisposed defendants cannot, even in these cases, avail themselves of the entrapment defense (although they may have some other defense available to them). Most courts have accepted the converse of this claim, as well; they have held, that is, that, assuming the defendant performed the act for which he is being tried as a result of a temptation provided by the government, the entrapment defense succeeds if it can be shown that the defendant was not predisposed.

One might understand appeals to predisposition as efforts to capture an intuitive distinction, the distinction between those who would have committed a crime like that in question in the ordinary course of things had the government not entered the scene, and those who would not have. Thus, we can formalize the Subjective Test as follows. Here, *S* is a defendant who performed an act of type *C*, a type of action banned by a particular criminal statute, and who did so as a result of a temptation supplied by the government:

The Subjective Test of Entrapment: *S*’s entrapment defense succeeds *if and only if* If the government had not performed any of the actions that

culminated in S's acceptance of the temptation it provided, it is not the case that S would have performed an act of type C on some other relevant occasion.

In most cases, this counterfactual "other occasion" will count as "relevant" if it might be encountered in the ordinary course of things even without special machinations on the part of the government. So, for instance in *United States v. Woo Wai* (223 F. 412 (1915 U. S. App.)), the defendant was given an opportunity by actual INS agents, whom he knew to be INS agents and who led him to believe they were corrupt, to smuggle illegal Chinese immigrants into the country. He evinced great reluctance to do so on the grounds that he would be caught. The agents assured him that he would not be caught and convinced him of this by noting that he would have the assistance of INS agents interested in concealing the crime. Since it is unlikely that Woo Wai would have, in the ordinary course of things, encountered corrupt government agents in position to aid in both the smuggling of immigrants and the concealing of the crime, and since there was ample evidence that he would only have committed a crime of the sort in question if it seemed extremely unlikely to him that he would be caught, it was ruled that there was not evidence sufficient to support the claim that he had the predisposition to commit the crime.¹¹

However, a possible occasion can be considered relevant even if there is good reason to think it very unlikely to occur without help from the government, if there is also evidence that the defendant was seeking it. For instance, in *Russell v. The United States* (411 U. S. 423 (1973)), an agent of the DEA supplied Russell with a substance that was very difficult to procure and that could be used to manufacture methamphetamine. Russell went on to use the substance for this purpose and sold the drug to the DEA agent, at which point he was arrested. The majority of the court held that the rarity of the substance was not relevant to the question of Russell's predisposition, since there was evidence that he was actively seeking to procure it and, in fact, had succeeded in procuring it in the past. In effect, the majority ruled that the occasion that would have prompted him to perform the crime — receipt of the rare substance — was relevant despite the fact that he was unlikely to have encountered it in the ordinary course of things. It is not hard to see the justification for this practice: A predisposition to *attempt* a crime is not less salient to responsibility than a predisposition to perform one. Perhaps neither is salient; but if the latter is salient, then so is the former. In cases like *Russell*, there is good reason to think that the defendant would have found an occasion to attempt the crime, for there is good reason to think that he would have found an occasion to make an effort to procure the rare substance, even if there isn't good evidence to suggest that he would have succeeded in that effort.¹²

Further, and relatedly, a possible occasion can be considered *irrelevant*, even if it's quite likely that the defendant would have encountered it in the ordinary course of things, for there might also be evidence that he was actively trying to avoid encountering it. In *Hampton v. United States* (425 U.S. 484 (1976)), for instance, the defendant, who had a past record of illegally procuring and distributing heroin, was attending a drug rehabilitation program in an effort to rid himself of his addiction. There he met an undercover government informant who, after repeated efforts, succeeded in cajoling him to acquire the drug so that the two could use together. The majority of the court ruled that Hampton was not predisposed, despite having had a track record of just such behavior in the past. Animating the decision was the thought that although the defendant might very well have fallen off the wagon in the ordinary course of things, the circumstances that would prompt him to do so were not "relevant" occasions for assessing his predisposition since he was actively seeking to avoid such occasions. The thought here is similar to that involved in *Russell*: Someone who is seeking to avoid the circumstances that will trigger illegal behavior on his part is best understood to lack a predisposition to attempt a crime, and so to lack the predispositions of a fully culpable defendant.

We can think of the Subjective Test for entrapment, then, as offering an account of predisposition. According to the test, a defendant is predisposed just in case, when we subtract consideration of the government's actions, we find that the agent still would have either performed, or attempted to perform, an act of the sort for which he is being tried. It is important to see that this is an account of *predisposition*, not an account of disposition generally. In some broad sense of the term "disposition," every defendant who gives in to the temptation to act as required for the crime is disposed to act that way: Part of what explains the fact that he did so act in response to the temptation provided is that he has a certain character or turn of mind, in short, a disposition. But not all such dispositions are predispositions in the sense of relevance here, because they are not all of them so sensitive as to prompt objectionable action on the defendant's part on "relevant" occasions, subtracting consideration of the government's behavior.

But why should predisposition, so understood, be thought to have any relevance whatsoever to culpability? And, further, why should the absence of predisposition constitute not just a mitigating factor, relevant, for instance, to sentencing, but, instead, a complete defense, a ground on the basis of which a defendant is excused entirely from responsibility? Advocates of the Subjective Test have answers to these questions available, although, as I will argue, the answers require acceptance of a retributive justification for criminal punishment over a deterrence or rehabilitation-based justification.¹³

Could the Subjective Test be justified through appeal to special deterrence — deterrence of the recipient of the punishment from criminal activ-

ity? As a general rule, one can justify allowing a defense to a particular criminal defendant on the grounds of special deterrence by pointing out that something about that defendant made it the case that the prospect of punishment could not, in fact, or could not fairly have been expected to deter him. So goes the standard special deterrence justification for not punishing those who kill in self-defense: Someone who believes he will be killed if he does not kill an attacker cannot be expected to be influenced by the prospect of punishment; whether or not he is to be punished, he probably will, and in any event still has the right to protect his life by killing the attacker. But notice that a parallel line of thought cannot hope to explain why the unpredisposed, who nonetheless give in to government temptation, are not to be punished; there is no reason to believe them to be incapable of resisting, nor is there reason to think that they had some special right to accept the temptation, even in the face of the prospect of punishment. This is especially clear given that, normally anyway, unpredisposed defendants believed themselves to be tempted by a private party, instead of the government; given that fact, what about them could make make them inapt targets of deterrent pressure?

How would an advocate of a general deterrence justification for criminal punishment — the deterrence of people generally from performance of certain acts by attaching punishment to them — explain the relevance of predisposition to legal culpability? Here the prospects for an adequate justification of the defense seem brighter. The policy of denying the defense to the predisposed and granting it to the unpredisposed is a fairly efficient way to decrease crime. Deterrent pressures are a societal cost; they should be exerted only if by doing so crime rates can be substantially reduced. In fact, there is reason to think that the policy enshrined in the entrapment defense under the Subjective Test succeeds in spending little of this currency in exchange for a substantial decrease in crime. To see this, start with an analogy: Imagine that there have been a rash of burglaries in a particular neighborhood and the police are considering adopting one of two ways of reducing the numbers. The first proposal is to have regular patrols through the neighborhood; the second proposal is to require anyone entering the neighborhood to prove himself to be either a resident or invited by a resident. Both proposed plans, we can be confident, would reduce the number of burglaries in the neighborhood, and let's assume that they would result in the same reductions. The first plan is obviously superior to the second, however, because the first involves exerting deterrent pressure only on those who are already planning to commit burglaries, while the second involves placing similar pressure on a large number of people who are not; by being asked to prove residency, every resident is reminded that he is being watched. By adopting the first policy over the second, we can achieve the same reductions in burglaries while exerting less unpleasant deterrent pressure. Similarly, we might say, by denying the entrapment defense to the predisposed, and grant-

ing it the unpredisposed, we will manage to reduce the rates of occurrence of certain crimes — particular crimes that are hard to detect in other ways like drug sales, prostitution and political corruption — while avoiding unnecessary deterrent pressures, pressures that won't serve to lower crime rates relative to the ordinary course of things. The predisposed, who would have acted criminally in the ordinary course of things, remain under pressure not to act criminally in a legal system that accepts the Subjective Test, while the unpredisposed, who would not have, do not. According to this line, that is, the Subjective Test is a way of ensuring that the government is employing the minimum amount of coercive pressure needed to bring about less crime than there would have been without government involvement. It assures, that is, that the government won't expend deterrent pressure on those who do not need to be deterred.¹⁴

The trouble with this approach is that it assumes, falsely, that no punishment need be threatened to deter the unpredisposed; it assumes, that is, that the unpredisposed simply won't act criminally if the government doesn't get involved, while all that follows from the concept of predisposition is that the unpredisposed are *unlikely* to do so. This is how the situation in entrapment differs from our hypothetical involving neighborhood burglaries. In the hypothetical, we are assuming that residents and those invited by residents simply won't commit burglaries, and so there is no point in exerting deterrent pressure on them. But the analogous assumption about the unpredisposed who are nonetheless disposed to accept certain temptations is false. The unpredisposed who give in to temptation are disposed to commit the crime; they differ from the predisposed only in that the circumstances that would prompt them to do so are peculiar or highly unlikely to be encountered. Thus, the view on offer does imply that given scarce resources the government can more effectively prevent crime by targeting the predisposed; they, after all, are more likely to commit crimes. But it doesn't imply that, once the government has succeeded in ensnaring an unpredisposed person that it ought not to punish him. On the contrary, to fail to do so is to fail to deter unpredisposed people who happen, as unlikely as it is, to find themselves, like the defendant, tempted to commit the crime. In short, a general deterrence approach can explain why the predisposed are denied the entrapment defense, but it can't explain why the defense is granted to the unpredisposed.

There's a rejoinder to be considered here. It might be suggested that while there is something lost in granting the entrapment defense to the unpredisposed, since doing so removes the deterrent pressure from those unpredisposed people who are still disposed to commit the crime, there is also something gained: The government is effectively discouraged from placing those who are not only unpredisposed but entirely undisposed under deterrent pressure. After all, the government can rarely, if ever, tell if an unpredisposed person is undisposed to commit the crime; usually, the way this is dis-

covered is by issuing the temptation and seeing if it is accepted. But the result of this behavior on the part of the government is that many undisposed people will suffer pressure from the government that they oughtn't have to endure; they will be, that is, much like the innocent residents of our imagined neighborhood who would be forced to justify their presence in their neighborhood to suspicious police, were the policy of screening people before entering the neighborhood adopted.

Notice, however, that this rejoinder shifts the ground in an important way. It involves abandoning the idea that the unpredisposed who accept government-issued temptation are not culpable and, instead, admits that they are culpable but insists that they should nonetheless not be punished so as to save the truly innocent undisposed people from unnecessary government prodding. Or, in other words, the rejoinder claims that what is really being accomplished by granting the entrapment defense to the unpredisposed is that a certain form of objectionable government conduct — the placing of pressure on the undisposed — is being discouraged. But this suggests that under the approach of the rejoinder under consideration, the entrapment defense is not substantive, but procedural: Those who are granted the defense are granted it because the government has behaved improperly toward them, and not because they are less than fully culpable for what they have done. This approach will be discussed in the next section of the paper. For now what's important is only this: At best, the Subjective Test can be justified through appeal to general deterrence only at the cost of its subjectivity, as it were; at best it can be justified through appeal to general deterrence only if the entrapment defense is, at bottom, about government misconduct and not about the defendant's diminished culpability. For our purposes in this section, such an approach gives up the game. The question at hand is what justification of punishment an advocate of the Subjective Test must accept, on the assumption that he takes the distinction between the unpredisposed and the predisposed to align with a distinction in culpability. What has emerged, and will be discussed in the next section, is that someone who takes the entrapment defense to be procedural rather than substantive might still think that lack of predisposition is sufficient for granting the defense.

So much, then, for deterrence. What are the prospects for justifying use of the Subjective Test from within a rehabilitation view of punishment? Someone who accepts a rehabilitation view must claim that the unpredisposed either can't be rehabilitated, or, more plausibly, are not in need of rehabilitation, while the predisposed, by contrast, both can be rehabilitated and need to be. At most, however, it seems easier to rehabilitate the predisposed, but unpredisposed, defendant than the predisposed and therefore there is at most a reason to give such defendants lighter sentences, and no reason to excuse them entirely. It might be argued, however, that under a rehabilitation-based view there is a threshold for objectionable disposition that is re-

quired for punishment. In order to view a defendant as in need of rehabilitation, for instance, we might need to think him *likely* to act objectionably again and not simply prone to do so under some unlikely circumstance, such as those that prompted the unpredisposed defendant to action. So construed, a rehabilitation view can seem to fit snugly with the Subjective Test: The Subjective Test draws the line, we might think, between those who are and those who are not prone to recidivism in the ordinary course of things, and thus in need of the supposed corrective influence of punishment. However, if predisposition walks in lock-step with a disposition to act similarly in the future, and if it is the latter that accounts for the appropriateness of punishment, then to excuse a defendant just in case he is predisposed is to prejudge the question of whether or not an unpredisposed defendant could benefit from the rehabilitative influence of punishment. Why not grant an entrapment defense just in case the defendant can show that he is not disposed to do it again and allow that the lack of predisposition provides some non-decise evidence for the absence of such a disposition? Someone who holds a rehabilitation view of punishment, that is, cannot also hold that the Subjective Test identifies the fundamental fact that separates successful from unsuccessful entrapment defenses, even if he thinks that predisposition is found regularly with the kind of dispositions on the part of the defendant, dispositions for future criminal conduct, that he takes to warrant punishment.

Earlier it was claimed that only someone who accepts a retributivist justification for legal punishment has the resources to justify adequately the Subjective Test for entrapment. So far, it has been shown that neither a deterrence- nor a rehabilitation-based view can explain the Subjective Test's emphasis on predisposition, as opposed to mere disposition to act wrongly. Thus it appears that an advocate of the Subjective Test must accept whatever justification for punishment remains, and this probably means accepting some form of retributivism. But why should it be thought that a retributivist view can do any better? Perhaps the approach embodied in the Subjective Test is simply inadequate for justifying the entrapment defense.

Retributivist views of punishment are notoriously vague. They say that legal punishment is justified just in case its recipient deserves it, but they are rarely coupled with anything more than a hand-waving attempt to explain what desert, in a legal context, really means. It is, of course, non-obvious what *moral* desert amounts to, but even if this is understood well-enough to play a robust role in a theory, it cannot help retributivists who (like myself) are drawn to broadly positivist accounts of law: If the content of the law is not stipulated or constrained by the content of morality, it is difficult to see what notion of "desert" is being invoked in the retributivist justification of punishment. In section 4, I'll offer a necessary condition for desert in the sense that's relevant to the law, but the point immediately at hand can be appreciated without having any such condition on the table: There are no prin-

ciplined pre-theoretical reasons to think that an account of legal desert could not be crafted under which defendants in cases of the sort at issue are deserving of acquittal on the grounds they offer just in case they were unpre-disposed. That is, without knowing more about what is involved in legal desert, there is no reason to think that a theory of it could not do the work that is needed. This is not a defense so much as a promise, a promise on which I'll make good before the end.¹⁵

Before ending this section, it is worth saying a word about what might seem to be a pressing objection to the Subjective Test. It might seem that the advocate of the Subjective Test, especially in cases in which the defendant's entrapment defense is rejected through employing the test, advocates holding people legally responsible not for conduct they commit, but for conduct that they *would have* committed in the ordinary course of things, in, that is, the counterfactual scenario in which the government did not provide the temptation to commit the crime. So put, this seems monstrous. It seems as bad, for instance, as simply detaining people for crimes they haven't committed and then convicting them for those crimes on the grounds that they would have committed them if not detained.¹⁶ But is the advocate of the Subjective Test committed to this? No, for those defendants whose entrapment defenses are denied on the grounds that they were predisposed, by the standards of the Subjective Test, are not being convicted for what they might have done, but for what they did in fact do. They are being claimed to be just as culpable for their actions as somebody who acted the same way without ever having been tempted by the government. The Subjective Test, that is, should be understood as a test for determining whether the acts the agent actually committed in response to government temptation arose in the normal way, the way required for legal culpability, or not. The question before the court in such cases, that is, is whether or not the defendant is legally culpable for the actions he actually performed; the test determines if this is so by assessing the truth of a counterfactual. That counterfactual is intended to tell us if what the defendant actually did in response to the government's temptation is reflective of the facts about the defendant that would justify punishing him.

Notice, however, that this answer to the objection reveals that the advocate of the Subjective Test is committed to the following result: The deliberative process that gave rise to the defendant's decision to accept the government's temptation must be a *continuation* of a past deliberative process, of which we have evidence, and which indicates that the agent would have performed the same actions had he found himself in a relevant circumstance. To illustrate, recall the case of Hampton, the drug addict who sold heroin to a government informant whom he had met in a drug rehabilitation program. One way of interpreting the case is this: The evidence suggesting that Hampton was predisposed was all evidence that predated his decision to enter a

drug rehabilitation program; hence, although it did indicate that at an earlier time Hampton would have sold heroin in a relevant circumstance, the features of Hampton that made that true are no longer present. To take this line is to deny that there is identity between the features of Hampton that support the claim that he was predisposed and the features of Hampton that gave rise to his decision to accept the government's temptation. But, there must be identity if predisposition is to be relevant to culpability not for what the defendant might have done, but for what he did in fact do in response to government temptation. We thus have discovered something about what those who take predisposition to be relevant to culpability are committed to: They must hold that evidence of predisposition tells us something about the very deliberative process that gave rise to the defendant's action, and not merely something about a distinct past deliberative process. This commitment will play some role in the solution to the problem of private entrapment developed in section 4.

The primary result of this section — the claim that advocates of the Subjective Test are committed to a retributivist justification of punishment — while not supplying a solution to the problem of private entrapment tells us what such a solution would have to look like if offered by an advocate of the Subjective Test: A solution would have to specify a necessary condition for desert of legal punishment, and it would have to show that just those whose defenses are granted under the Subjective Test (the governmentally tempted unpredisposed) fail to satisfy that condition, while just those whose defenses are denied under it (the privately tempted and the governmentally tempted predisposed) do satisfy the condition. While I will undertake to provide just such a solution in section 4, it can appear a daunting task, daunting enough that we might hope to find an easier solution to the problem of private entrapment in the Objective Test and its associated justification for the defense.¹⁷

3. The Objective Test and the Re-Appearence of the Problem of Private Entrapment

Impressed by the problem of private entrapment, we might insist that the real justification for excusing the entrapped is not the innocence of the unpredisposed, but the objectionable nature of certain forms of treatment by the government. Thus, according to any version of the Objective Test, a defendant can successfully employ an entrapment defense just in case he can show that the government overstepped its bounds in providing him with a temptation to commit a crime. The difficult task for someone who endorses some version of the Objective Test is to specify the precise conditions under which the government's action has tainted the process through which the defendant is arrested and tried, and will be punished if his defense is denied.

Different forms of the Objective Test will specify these conditions differently.

It helps to have a particular version of the Objective Test to work with, a version, that is, that specifies precisely what the government has done wrong in cases in which an entrapment defense should be honored.¹⁸ According to the *Model Penal Code's* version of the Objective Test, a defendant is entrapped when the government has employed

methods of persuasion or inducement that create a substantial risk that such an offense will be committed by persons other than those who are ready to commit it. (*Model Penal Code* §2.13(1)(b))

This test specifies the nature of the impropriety of government conduct in entrapment cases not by describing that conduct, but, instead, by describing its effect: It creates a risk of ensnaring someone who was not predisposed to commit the crime in question. The test draws the lines differently from the Subjective Test because it, unlike the Subjective Test, excuses those who are predisposed but accepted an objectionable government-issued temptation, objectionable because it risked ensnaring the unpredisposed.¹⁹ The *Model Penal Code's* test formalizes an idea that we saw, already, in discussion of the prospects for a general deterrence justification of the Subjective Test in the previous section. The idea behind the *Model Penal Code's* test is this: What the government does wrong in cases in which the entrapment defense is to be honored is to exert deterrent pressure on those who do not need to be deterred from committing crime, namely the unpredisposed; thus, according to the *Model Penal Code's* test, the government oversteps its bounds by using more coercive pressure to prevent crime than is actually required for the task. To put the point even more simply, the government should only be authorized to engage in behavior that causes criminal conduct if, by doing so, comparable or worse criminal conduct is also thereby prevented; by engaging in behavior that risks ensnaring the unpredisposed, however, the government engages in behavior that is likely to cause criminal conduct without preventing any, since the unpredisposed are not likely to have acted criminally without temptation from the government. The *Model Penal Code's* test, then, specifies the objectionable nature of government conduct in cases in which the entrapment defense is to be honored by focusing not on what the government does, *per se*, but on who they do it to.

The difference between the *Model Penal Code's* test and the Subjective Test can be illustrated by comparing a pair of cases: *Jacobson v. United States* (503 U. S. 540 (1992)), the most recent Supreme Court case on entrapment, and *United States v. Gendron* (18 F.3d 955 (1994 U.S. App.)). Jacobson was a farmer who had mail-ordered a pornographic magazine that included depictions both of naked children and adults in the early 1980s, before child por-

nography was made illegal. In the late 1980s, after laws against child pornography had been passed, the FBI got Jacobson's name and address after seizing the records of the bookstore that sold him the magazine. Over the course of the next two-and-a-half years, the FBI sent him volumes of fake catalogs, from various fictitious companies, advertising child pornography, as well as questionnaires from fake companies intended to assess his interest in child pornography. He responded to some of the questionnaires and indicated that he was interested in naked pictures of boys in their mid-to-late teens, a category defined broadly enough to include materials that he had every legal right to order. After receiving one catalog from a fake company that claimed to fund its free-speech lobbying efforts with the proceeds from the sales of pornographic magazines depicting children, and which declared the government's banning of these materials "hysterical nonsense," Jacobson ordered a magazine. He later claimed that he was interested to see for himself if the government's ban of these materials really was hysterical nonsense; he even claimed that he was uncertain if the materials would, in fact, depict minors. On delivery, he was arrested for receiving child pornography through the mail. The majority in the case endorsed the Subjective Test, and honored Jacobson's entrapment defense on the grounds that Jacobson was not predisposed to commit the crime in question.²⁰ In fact, an advocate of the *Model Penal Code's* test is likely to have reached the same verdict: What the government did to Jacobson risked ensnaring an unpredisposed person. The mere fact that the government did, in fact, ensnare an unpredisposed person is not sufficient to show that its behavior risked that event — often a harm will be caused by behavior that doesn't risk that harm — but, still, in this case the government's behavior does indeed seem to risk ensnaring unpredisposed people and so the advocate of the *Model Penal Code's* test will want to honor Jacobson's entrapment defense.

Gendron, like Jacobson, had bought child pornography before it was illegal and, like Jacobson, was targeted by the FBI, which sent him numerous catalogs from fake companies advertising child pornography. In fact, he received many of the same fake catalogs as Jacobson, including the catalog from the company claiming to fund free-speech lobbying with the proceeds from sales. He also used the entrapment defense. There are, however, important differences between the two cases, differences that might seem to make a difference. Gendron, unlike Jacobson, evinced no reluctance about accepting the government's offers but, rather, pursued them zealously. Further, Gendron, again unlike Jacobson, seems to have had entirely prurient motives for buying child pornography: On one of the orders that he places he describes the materials being advertised as those he has been "dreaming of possessing," and wonders if there is a discount available to those willing to buy in bulk (Appendix to *United States v. Gendron* (18 F.3d 955 (1994))). The difference between Gendron and Jacobson is just this: Gendron is predisposed

and Jacobson is not. In fact, in this case the court again endorsed the Subjective Test and, judging that Gendron, in contrast to Jacobson, was predisposed to commit the crime, denied the entrapment defense, convicting Gendron. By contrast, an advocate of the *Model Penal Code's* test would have acquitted Gendron. After all, assuming that what the government did to Jacobson risked ensnaring the unpredisposed, and given that it did the same things to Gendron as to Jacobson, it follows that its conduct was just as bad in the two cases.

Some will see this result, all by itself, as pointing out a flaw in the *Model Penal Code's* test.²¹ After all, we might say, aren't the Gendrons of the world — people preparing to commit crimes behind closed doors that are no less dangerous and damaging for our difficulty in detecting them — precisely the sorts of people whom we hope to catch using sting operations and other similar policing techniques? This, however, is far too fast, as anyone who thinks there should be *any* procedural defenses will see. Procedural defenses are always available to the culpable, and so any procedural defense will imply the possibility of excusing from punishment some people who are culpable for committing crimes. A coerced confession is no more admissible as evidence for the fact that with its help we are able to put a culpable offender behind bars. This is a result with which we must live if we are to impose constraints on what our government can do in the course of its efforts to punish crime.

To see the real problem with the *Model Penal Code's* proposed test, start with the following general point: If creating a risk of X is a bad thing to do, it is because X is a bad thing to actually happen. Driving at excessive speeds without hurting anyone is bad because of the risk it imposes of hurting someone; it wouldn't be bad if there were nothing bad about the actual occurrence of what it risks. So, if there is something bad about *risking* ensnaring the unpredisposed, it must be because there is something bad about *actually* ensnaring the unpredisposed. Conversely, if a defendant's entrapment defense is to be denied when the government's conduct does not risk ensnaring the unpredisposed, despite the fact that it obviously risks ensnaring the merely disposed, it follows that there is nothing bad, or nothing bad enough to discourage the government from doing it. Thus, a full justification of the *Model Penal Code's* test is going to require explaining why ensnaring the unpredisposed is a bad thing to do, and ensnaring the predisposed is not. It might be suggested, along the lines of the general deterrence approach discussed in the previous section, that when the government ensnares an unpredisposed defendant it acts inefficiently in its efforts to prevent crime by using resources that would have been more efficiently used if directed toward the predisposed. And this might be true. But has the government acted *so* inefficiently as to deny conviction of the unpredisposed? After all, the government's efforts have served to identify a disposed defendant, even if un-

predisposed. Thus, it has served to prevent certain unlikely crimes. To grant the defense in such a case, then, is to undo whatever good was done by the government's efforts simply because those efforts could have been better used. Economic-style arguments might be offered for this behavior — sometimes it's better to waste a dollar so as to prevent inefficient usages of future dollars than it is to reap whatever slight benefits it could provide. But conclusions reached through arguments of this sort must be limited by rights to equal treatment, rights that are trampled in this case. Any particular predisposed defendant has the right to an explanation for why he is convicted while the disposed, but unpredisposed defendant charged with the same crime is released. To say to this defendant that in his case, and not the other, the police used its power efficiently is to offer him the wrong sort of explanation for the difference in treatment. With such an explanation, the appeal of the Objective Test is lost. To say to our predisposed defendant that the government acted *outrageously* toward the unpredisposed defendant and not toward himself is to give the sort of explanation that would justify a difference in treatment. But there is no sense of outrage, or outrage of the needed sort, to be felt toward a government that merely acted inefficiently in one case and not in another.

The question, then, is how the advocate of the *Model Penal Code's* test can explain why it is outrageous for the government to ensnare an unpredisposed person while it is not comparably outrageous for them to ensnare a predisposed. Notice how attractive it is to say, at this point, that the reason it's bad to ensnare the unpredisposed is that an unpredisposed person is not really legally culpable for his behavior and, conversely, the ensnared predisposed person is culpable. However, the defender of the *Model Penal Code* can't offer this explanation without giving up the quick solution to the problem of private entrapment because the next question will be why it is that an unpredisposed person who is tempted by a private party is not excused also on the grounds that his culpability has been diminished. That is, this form of defense of the *Model Penal Code* simply allows the problem of private entrapment to return in a slightly different place.^{22 23}

Could a different version of the Objective Test from that offered by the *Model Penal Code* succeed? Acts of tempting can differ from one another intrinsically: Some are more insistent than others, for instance. Could intrinsic differences of this sort help to elucidate the distinction between objectionable and acceptable government conduct in a way that would still provide for a solution to the problem of private entrapment? Imagine that two acts of temptation by the government were aimed at qualitatively identical defendants, both of whom accepted the temptation, but the acts of tempting differed from one another; in the first case, say, the government was very insistent and in the second it was not. If we are to allow the first defendant's entrapment defense and deny the second's on the grounds of this difference

between the two acts of temptation, we must offer an explanation for why the intrinsic difference between the two acts (the difference in degree of insistence on the part of the government) justifies a difference in treatment of the two defendants. If we take the difference between the two government acts to be significant it must be because the one (the insistent temptation) is causing or risking a certain kind of harm that the other (the less insistent temptation) is not. But notice that we can't specify the nature of this harm through appeal to a difference in culpability between the defendants without returning to a form of the Objective Test that, like the *Model Penal Code's* version, reintroduces the problem of private entrapment. But what else could be normatively significant about the difference between the two acts of tempting? We might identify a right that is violated by the one act of temptation, and not by the other: The insistent act, for instance, might place the defendant under duress. But, in that case, the problem of private entrapment fails to be solved for a different reason: Duress induced by private parties excuses for just the same reasons that duress induced by the government does. If the differences between the two imagined acts of government temptation are too large, they will be present also in parallel cases of private temptation; but if they are not large enough, they will not justify a difference in treatment of their respective defendants. It seems that intrinsic differences between acts of government temptation matter to the degree to which they make a difference to *the culpability* of those who accept them. However, with this answer, irresistible as it is, comes the need for a solution to the problem of private entrapment. So far, then, it seems that the Objective Test loses its easy solution to the problem of private entrapment once the line between acceptable and unacceptable government conduct is specified adequately.

There is another way, however, to specify the objectionable nature of the government's conduct, a way that appeals neither to the characteristics of the temptation's target (as on the *Model Penal Code's* test) nor to the intrinsic features of the act of temptation. According to this approach, the standard by which to judge the government's conduct derives from the principles that guide the government's conviction and punishment of defendants, the reasons, that is, for thinking those activities to be legitimate. The government overstepped its bounds in issuing the temptation, on this view, just in case the issuance of the temptation involved behavior that is somehow in conflict with those principles. This is to think of the entrapment defense as a form of estoppel.²⁴ The term "estoppel," of course, means a variety of different things in the law, but in the sense that it is relevant here, a party, or the government, is estopped from doing something that it could only do given its prior commitment not to. In fact, there is a defense sometimes called "entrapment by estoppel," which is different from the standard entrapment defense. In cases in which entrapment by estoppel is a legitimate defense, the government has falsely convinced the defendant that the conduct for which

he is being prosecuted is, in fact, legal.²⁵ So, the thought goes, to allow the conviction would be to allow the government to express, by convicting and punishing, its commitment to the illegality of the relevant conduct when conviction and punishment are available to the government only because it earlier expressed its commitment to the denial of that. In other words, by convincing the defendant that the conduct was legal, the government committed itself to not convicting or punishing the defendant for that conduct; and since the defendant only acted as he did because of that commitment, to convict and punish would be to authorize governmental behavior while at the same time authorizing the commitment not to engage in it. This would be to involve the judge who convicts in a kind of contradiction or incoherence in which both the commitment not to do something and the doing of it are legitimized. Thus, we might say, the paradigm case of governmental misconduct is that involved in entrapment by estoppel, but what makes such conduct objectionable is that, were the conviction allowed, the government would be involved in a kind of incoherence.

Notice that estoppel of the sort involved in the defense of entrapment by estoppel only makes sense when the agent who is estopped from acting is the same one who was earlier committed not to so acting. There's no incoherence involved in authorizing A's commitment not to perform an act as well as authorizing B's performance of it. The incoherence only arises when A and B are the same agent. So there is no problem of private entrapment by estoppel. There's no mystery as to why conviction is allowed when a private party deceives the defendant into believing that a particular course of conduct is legal. The mistake on the defendant's part does not ameliorate his responsibility all by itself (we can assume²⁶), and there is no reason, in this case, to estop the conviction and punishment since the conviction and punishment are not actions on the part of the very same agent who is committed to the legality of the conduct. Thus, if the standard entrapment defense is a form of estoppel, then the problem of private entrapment is solved.

But is it? The case for saying "yes" might be put like this: The government is granted the power to convict and punish only so as to prevent crime. When the entrapment defense is to be granted, the government plays a causal role in producing the defendant's conduct. By causing the conduct for which the defendant is being prosecuted, however, the government commits itself to the position that this is not the sort of behavior that needs to be prevented; but to authorize the conviction and punishment is to say that the defendant's behavior is of that sort. Hence to authorize the conviction is to involve the judge in an incoherence; it is to say, at once, that the defendant's conduct is and is not of the sort that ought to be prevented. So, the argument goes, by granting an entrapment defense, the judge is saying that the government is estopped from convicting.²⁷ Since there is no similar

bar to authorizing the conviction when a private party does the tempting, the problem of private entrapment is solved.

As compelling as this case seems, however, it fails. To see this, notice that under this construal of the entrapment defense, the government is thought to commit itself to some sort of approval of the defendant's conduct by virtue of having tempted the defendant to engage in it. Tempting, in standard entrapment cases, is supposed to be analogous to convincing the defendant that the relevant conduct is legal, as in entrapment-by-estoppel cases. But why should we think that tempting someone to do something involves any such commitment? Or, to refine the point, say it is granted that tempting someone to do something involves an at least tacit form of approval of the conduct in question. Does it follow that the kind of approval in question is of the sort that the government would be denying or repudiating in convicting and punishing the defendant? No. By convicting and punishing, the government is committed only to this much disapproval of the defendant's conduct: It is illegal and, correlatively, of the sort that the government is authorized to punish (perhaps because the conduct is deserving of punishment). But if in tempting him the government never led the defendant to believe that the conduct it tempted him to perform was legal, or somehow undeserving of punishment, then it did not ever manifest a form of approval of that conduct that was in conflict with its later conviction and punishment of him. What follows is that the problem of private entrapment can't be solved by emphasizing a similarity between standard entrapment and entrapment by estoppel. The problem is that the defense of entrapment by estoppel derives its legitimacy from the fact that the government undermines the principles behind its own conviction and punishment by the actions that it uses to induce the defendant to act. But since simple temptations do not similarly undermine the principles behind conviction and punishment, the solution to the problem of private entrapment to which the entrapment-by-estoppel defense can appeal is not available to the standard entrapment defense. To put the point another way, in entrapment by estoppel, the principles that justify conviction and punishment can help us to identify a form of government conduct that amounts to misconduct; the government can't act in a way that involves commitment to the legality of the defendant's conduct and later premise its conviction on that conduct's illegality. But the issuance of temptations of the sort that we do and should take to amount to entrapment do not necessarily conflict in a similar way with the principles guiding the government in conviction and punishment. Entrapment by estoppel is to be justified on "objective" lines, but the relevant form of justification cannot stretch to standard entrapment and still hope to quickly solve the problem of private entrapment.

It seems, then, that there are three broad forms of the Objective Test, or three ways of specifying the line between acceptable and unacceptable

able government temptations. The government can be thought to act wrongly by virtue of (1) who it targets or risks ensnaring, (2) the intrinsic features of its acts of tempting, or (3) conflict between the principles to which it is committed in issuing the temptation and those to which it is committed in convicting and punishing. Versions of the Objective Test of the third variety, it has been argued, overlook the fact that many temptations that amount to entrapment do not involve principled commitments that are in conflict with those guiding conviction and punishment. More importantly, this section has argued that versions of the Objective Test of either the first or second variety re-encounter the problem of private entrapment: They may be viable, but only when supplemented with an explanation for how the culpability of those whose entrapment defenses are to be honored has been undermined by the government's conduct. Advocates of the Objective Test, then, do not earn, thereby, an easy solution to the problem of private entrapment.

4. The Solution to the Problem of Private Entrapment

At the end of section 2, an abstract description was given of the kind of solution to the problem of private entrapment required by an advocate of the Subjective Test: A solution must specify a necessary condition for desert of punishment and must show that the unpredisposed and privately tempted fail to satisfy it while the predisposed and governmentally tempted do not. The results of section 3 place us in a position to see that a very similar sort of solution to the problem of private entrapment is required by the advocate of an adequate version of the Objective Test: A solution must specify a difference in culpability between the unpredisposed and privately tempted, on the one hand, and the governmentally tempted and predisposed, on the other. Since, as was argued in section 2, such a difference in culpability can't be found without appeal to a retributivist justification of punishment, drawing the needed line requires specifying (at least) a necessary condition for desert unsatisfied by the former set of defendants, and satisfied by the latter. First, then, whether it is the Subjective or Objective test we advocate, we need a necessary condition for desert of legal punishment.

Start with the following form of informal justification for the issuing of a punishment, "He brought it on himself." When we say things of this sort to justify inflicting a person with a punishment, what do we mean? We don't take ourselves to be merely pointing out what is already known, namely that the defendant did something that deserves punishment. We seem to be saying more than we would were we to say instead nothing more than, "He deserves it." But what more is being said? We don't mean what we literally say, for we still think it is necessary for us, or the prison guards, to actually issue the punishment; we don't think, for instance, that left to his own devices the defendant who is to be punished will deprive himself of liberty, or

voluntarily pay a fine, even if given no incentive by the state to do so. Criminals almost never actually choose to be punished; they choose illegal acts either disregarding the punishment, or perhaps not thinking about it at all, or believing themselves to be capable of avoiding it. In what sense do the deservingly punished, then, "bring the punishment on themselves"?

To give some systematic content to this thought, it helps to have in hand a particular conceptual tool: the idea of pressures on deliberation, particularly pressures regarding what features of possible actions (including, although not limited to, their consequences) we grant reason-giving force in deliberation. Deliberation issues in decision and the formation of intention following reflection on various possible actions, and the weighing of reasons for and against the actions considered. But, in deliberating, an agent faces a wide variety of normative constraints: He is under pressure with regard to the process through which a decision arises; if the decision arises in the wrong way, then his deliberation has fallen short of what it ought to be. Some of these constraints have their source in prudential rationality: An agent is under peculiarly rational pressures with regard to deliberation — pressures, for instance, to consider actions that he believes to be means to actions he intends. Among the wide range of rational constraints on deliberation are constraints regarding what features of considered actions are and are not to be granted reason-giving force. Rationality prescribes that some possible causal consequences, for instance, need to be considered and weighed. But which ones? Just the likely consequences? How likely is likely enough? Just the consequences that are to befall the deliberator? Or must he also consider consequences of his action for others? If so, which others? Just those whom he cares about? Just those who can help him to get what he wants? In debating the answers to these questions we debate about what prudential rationality requires of deliberation.

Some pressures on deliberation, even with regard to what features of an action to consider, are peculiarly moral: Morality might require, for instance, that one grant reason-giving weight to other people's pain, even those whom one hates, and might require that one give greater weight to such pain than the pain of animals, for instance. It seems plausible that among the consequences of possible actions that morality requires us to reflect upon are those consequences that would come about were the action to be rewarded with everything that it deserves morally. Agents labor under pressure to grant reason-giving force to what would happen were they to perform a particular act in a world in which all transgressions are met with their just desserts; and they are under such pressure regardless of their assessments of the likelihood that this will happen. If conduct you are considering would *warrant* your becoming a pariah, then you are under pressure to weigh that consequence in your deliberations even if you are all but certain that you won't become a pariah even if you act that way. In addition, and of particular importance, is the

fact that agents labor under pressure to grant reason-giving weight to the mere fact that an act is moral or immoral, independently of the consequences that are attached, in a perfectly just world, to moral or immoral action. If an act is immoral, agents labor under pressure to grant that fact itself reason-giving weight in their deliberations, independently of the consequences for themselves or others of performing such an act.

Agents labor under very similar pressure with regard to deliberation in so far as they are citizens of a state: For each action they consider, they are under pressure to seriously consider what will happen if the government responds to that action in all the ways the law requires and allows the government to respond, and they are under pressure to take seriously the legal status of each considered act — its legality or illegality. Agents labor under pressure, that is, to grant reason-giving weight to the *legal consequences of action*, consequences that may or may not actually come to pass, but that would come to pass were the law to function perfectly; and, more importantly for our purposes, they labor under pressure to grant reason-giving weight to the consequence-independent legal status of the acts they consider.

In addition, that an agent is under pressure to grant a particular legal consequence or the legal status of an act some reason-giving force in his deliberations does not imply that *he knows* he is under such pressure. Say, for instance, that it is illegal for me to spit on the sidewalk, but I don't know that it is. One of the legal consequences of spitting on the sidewalk, say, is a public caning. I am, therefore, under pressure to grant the caning reason-giving force in my deliberations about whether or not to spit on the sidewalk and I am under similar pressure to grant the illegality of spitting on the sidewalk reason-giving force. My deliberations have failed to live up to the standards imposed by the law when I fail to grant these features any reason-giving force, despite the fact that I have no idea that the act is illegal or that such a punishment is mandated by the law. Notice that we might be tempted to say that I am under legal pressure to grant these features reason-giving force because I am under legal pressure to inform myself about what the law requires of me. But, in fact, this oversimplifies the situation. The law doesn't require that we know what it requires; it only requires that we conform. The obligation to conform requires us to overcome various obstacles to conformity, including ignorance about what the law requires, but overcoming ignorance is required of us only in so far as failure to do so would result in non-conformity. What this shows is that the concept of legal pressures on deliberation is not a psychological concept, but a normative one. A person who is under such pressure may or may not feel it, and may or may not understand it. It is possible that in order to be under certain pressures in deliberation, you need to have had fair opportunity to conform — perhaps this is the root justification for the requirement of publicity of the law in a just legal system — but, nonetheless, we cannot hope to determine whether or not an agent

was under pressures in his deliberations merely from examination of his psychological and cognitive states.

What is the difference between saying "S was under pressure to grant such-and-such legal features of C weight in his deliberations" and saying, merely, "C-ing is illegal and will be given such-and-such a punishment"? After all, we might say, if the first statement is true, the second is, and vice versa; doesn't that show that there is really no difference between them? The difference between the two statements can be seen by noting that a law specifying punishment for an act imposes two norms, not just one: It imposes the obvious norm on conduct; it says you are doing the wrong thing, from a legal point of view, if you act as described in the law. But it also imposes a distinct norm on deliberation; it implies that you are deliberating in a less than fully adequate way if you fail to grant the legal features of the act reason-giving weight in deliberation. Of course, it is only violation of the first of these norms, and not the second by itself, that ever warrants punishment; but still, it is a fact, and an easily overlooked one, that the second norm is imposed also when an act is made criminal.

With the notion of pressures on deliberation in hand, return to the question of what we mean when we say of a person who is being punished as the law specifies, "He brought it on himself." What we mean stems from the observation that an agent who engaged in fully informed and procedurally perfect deliberation, and who was considering the act he performed, would have granted reason-giving force to the illegality of the act. Thus, someone who performed an illegal act without giving its legal status any thought, or without granting that status any weight in his decision about what to do, didn't deliberate consonant with the full range of pressures on deliberation; and somebody who did take the act's illegality into consideration, and chose the act anyway, chose the act while allowing the illegality to figure into his deliberations as a reason not to perform the act, but a reason that was outweighed by other goods the act promised. He, thus, chose the illegality of the act itself in the following weak sense of the term "choose": *The reason-giving force of the act's illegality was included in the calculus of reasons that entered into his deliberation and he nonetheless chose an action that was illegal.* Since the punishment is appropriate just in case the act was illegal, when an agent chooses (in this weak sense) an illegal act, he also chooses (in the same weak sense) the mandated punishment. It would be, therefore, too strong to say that the defendant incurs the punishment *voluntarily*; the sense in which the punishment is chosen is too weak for that. But, still, incurring the punishment has a feature in common with consequences brought about voluntarily: It is the upshot of a feature of the act the reason-giving force of which is weighed, or ought to have been. This second disjunct is particularly important: A person can have brought a punishment on himself even if he never gave the illegality of the act any thought and even if he had no idea that the act was illegal, for the

absence of these psychological states does not necessarily undermine the fact that he is under pressure in deliberation to grant the act's illegality reason-giving weight.

This much explains why it is that a person does not deserve punishment, as specified by a statute, for an act performed before the statute banning it was made law. Such a person could not be said to bring the punishment on himself, for at the time of action the agent did not labor under any pressure to grant the illegality of the act reason-giving weight; he could not have since the act was not, at the time, illegal. But if he was under no such pressure, then he can't be said to have, even in the weak sense, chosen the punishment, and so does not deserve it. The sense in which the legal pressures on deliberation are distinct from psychological pressures emerges here, also, although in a different way. Imagine that a person performs an act believing it to be illegal despite the fact that there is no statute banning it. Now imagine that, later, a statute banning it is made law and the statute is retrospective: It says that acts of the relevant sort were illegal before the statute was made law.²⁸ Notice that if the pressures on deliberation were psychological, then there would be no reason to think it unjust to punish the person in our example: After all, he believed himself to be choosing an illegal act and so believed himself to be under legal pressures in deliberation that he elected to ignore. However, it would be unjust to punish him and for just the reasons being described here: He was not, in fact, under pressure to grant the illegality of the act reason-giving weight, even though he thought he was, and so cannot be said to have chosen the punishment in the weak sense and in the face of such pressures.

However, more can be said about what is required to have brought a punishment on oneself. To see this, note first what I will call "the canceling effect," a phenomenon most easily illustrated from consideration of the prudential pressures on deliberation, pressures to grant various causal consequences of action reason-giving weight. Imagine that all of the acts one is considering (and this list includes omissions of acts) will result in the occurrence of a particular result; that result, from a practical point of view, is inevitable. Are you under prudential pressure to grant the result reason-giving weight in your deliberations about what to do? No, for the granting it weight in your deliberations cannot have any effect on the outcome of those deliberations. It adds equally, in the calculus of reasons, to each of the acts you are considering and so cannot make it the case that any one of them is favored more strongly by reasons than the others. This is the canceling effect: The inevitability of a feature of action — its being a feature of all the different actions an agent is considering — cancels the need to grant it any reason-giving weight in deliberation. This doesn't mean that the agent *won't* grant it any weight — he might still; it means only that he hasn't failed to live up to the standards of perfect deliberation if he does not. Another way to put the

point: When a feature of action, such as a consequence of action, is subject to the canceling effect — each of the various possible actions the agent is considering possess it — the agent's failure to grant it any reason-giving weight in deliberation is not a failure on his part to exercise correctly the peculiar, and peculiarly human, capacities that are required to engage in deliberation.

Normally, an agent cannot be said to have brought a punishment on himself if the illegality of the act for which he is punished was subject to the canceling effect. Imagine, for instance, a state in which the law says simply that anyone who knowingly causes, or knowingly allows, the death of a person will be punished, and specifically includes both cases of killing in self-defense and suicides among those killings that it bans. A person who finds himself having either to kill another who threatens him or to allow himself to die at that person's hand, then, is facing a situation in which all of his alternative actions are illegal; he must either knowingly kill another, or knowingly allow his own death, and both acts are banned by the law. Such a person can't be said to have brought the state's punishment on himself, even if he did, in fact, grant reason-giving weight to the illegality of each possible act in his decision about what to do: Everything that he was considering doing was illegal. Thus, whether he granted the illegality any weight in his deliberations, or none at all, could not have had an impact on the deliberation's results. Since the illegality of his various alternative actions was subject to the canceling effect, in this case, he was not under normative pressure to grant it reason-giving weight, and so can't be said to have brought the punishment on himself when he is punished. Such a state, then, has a flawed criminal law: It requires punishment of the undeserving.

Notice, however, that this conclusion would be undermined were we to elaborate the story a bit: Imagine, for instance, that the man planned to kill this other person and took steps to do so. In the process, the other person attacked him and made it clear that he would kill him if not killed trying. Now, although the man in our example still faces a situation in which every action he is in position to consider is illegal, he can still be said to have brought the punishment for killing on himself. The additional piece of information — that the man was plotting the death of his assailant — undermines the excusing force deriving from the inevitability of performing an illegal act. Here, it seems, the man was under pressure to grant the illegality of the act reason-giving weight *prior* to the situation developing into one in which the illegality became subject to the canceling effect. Hence, since there was a time at which the illegality was or ought to have been granted meaningful reason-giving weight, the punishment was, in fact, chosen in the weak sense; the man can still be said, in this instance, to have brought the punishment on himself.

We thus have the following necessary condition for a punishment to be deserved:

S deserves to be punished for C *only if* At some prior time, S was under legal pressure to grant the illegality of C reason-giving weight in his deliberations and decided to do C, despite those pressures.

As we've seen, an agent might not deserve to be punished because in all of those deliberations in which he considered C-ing, the illegality of C was subject to the canceling effect. And, an agent can deserve to be punished, even if the illegality of C was subject to the canceling effect in some of his deliberations about C so long as in other deliberations C's illegality was not subject to the canceling effect and he chose C. In order for punishment to be deserved, we need to find some past deliberation in which the agent was genuinely under legal pressure to consider the act's illegality, to grant it reason-giving weight in his deliberations, and chose an act of the relevant sort in the face of such pressures.

A number of further complexities require note: First, notice that there is no canceling effect when the features of alternative actions are of different sorts; legal consequences do not cancel the pressures of prudential consequences, nor do moral cancel legal, for instance. Imagine that your boss threatens to dock your pay by \$1,000 if you refuse to be involved in a form of accounting fraud the penalty for which is a fine of \$1,000. Thus, assuming that your boss will certainly manage to dock your pay if you don't comply and that you will certainly be convicted of fraud if you do, compliance has a \$1,000 loss as a legal consequence, while refusal has a \$1,000 loss as a causal consequence. It would be silly to say that you are under no pressure to grant the punishment reason-giving weight in your deliberations about what to do, or that you haven't brought the punishment on yourself when you play ball with the boss.²⁹ But nor is the view on offer here committed to these silly results, for there is no reason to expect causal consequences to cancel legal, or vice versa. Similarly, there is no reason to expect the illegality of one possible action to cancel the immorality of another, or vice versa. Deep questions linger here: How are we required to deliberate when there are reasons of different sorts favoring competing actions? But hard questions of this sort are best understood as questions about the conditions under which reasons of different sorts *trump* one another and not as questions about the conditions under which reasons of different sorts *cancel* one another; canceling is a phenomenon confined to single kinds of reasons.³⁰

Another complexity: The canceling effect is a matter of degree, and in two different ways. First, there can be some degree of canceling when different considered alternative actions have consequences of different values. Say, for instance, that you are trying to decide whether or not to do act A or

act B. A will result in a net loss of \$100 while B will result in no net loss of funds. If you choose A, thus deciding that it is worth the \$100 cost, then you choose that loss in the weak sense of choice: You are under pressure to grant the loss reason-giving force in your deliberations and yet you elect a course of conduct that has it as a consequence. But now imagine that, instead of resulting in no net loss, B also would result in a \$100 loss. Now the \$100 loss is subject to the canceling effect: You can just ignore it in your deliberations and you won't be deliberating in opposition to the pressures of prudential rationality; maybe you give it weight, but you are under no pressure to do so. But it would be somewhat subject to the canceling effect, although not completely subject, were the loss for performing B \$40 instead of \$100. Then the \$100 loss associated with A need only be granted the reason-giving force of a \$60 loss. It is somewhat, although not entirely, canceled by the presence of a comparable evil's attachment to the other acts being considered. This last is an example of the first way in which the canceling effect can be a matter of degree.³¹ Similarly, canceling can be a matter of degree even when the features of the alternative actions are consequence-independent. If you are choosing between committing a burglary and committing a murder, the illegality of the burglary does not entirely cancel the illegality of the murder. Assuming that murder is strictly worse, legally speaking, than burglary, you remain under pressure to grant the illegality of the murder reason-giving weight, but do not labor under pressure to grant the illegality of the burglary any weight.

Before moving to the second way in which canceling is a matter of degree, notice that when we complicate our picture by considering features of action the values of which are not easily, or only clumsily, expressed in monetary terms, the question of how much canceling there is is the question, which won't be answered here, of what is required of us in the face of incommensurable, or only partially commensurable values. It would be not just silly, but deeply distorting, to suggest that a person deciding whether or not to accept \$100 in exchange for dancing on the table might be under no pressure to consider the loss of dignity associated with accepting; it would be obviously and deeply mistaken, that is, to suggest that under certain circumstances (those in which the disutility of the indignity is just the same as that associated with, say, a \$90 loss), there is no prudential pressure to grant the indignity reason-giving force. The disvalue of indignity is not the sort of thing that can necessarily be monetized. Still, it would be easy to overstate the point. The choice between a monetary loss and a loss of one's dignity (the choice put by the offer to dance on the table) is harder than the choice between a lesser monetary loss and the loss of one's dignity (imagine that the offer was for only \$50). But why? The difference in difficulty derives from the fact that a monetary loss does indeed cancel the disvalue of the indignity to some degree, and thus a lesser monetary loss cancels it more, even if it is

difficult, or most likely impossible, to quantify the degree of canceling involved. A loss of an apple has an ameliorating effect on the gain of an orange, even if they are, in the end, apples and oranges. In short, the first sense in which the canceling effect is a matter of degree is intended to capture, also, degrees of canceling involved in the comparison of features of action the values of which are incommensurable, or only partially commensurable.

However, there is another way in which the canceling effect is a matter of degree that will turn out to be more important for our purposes: A particular feature of action can be more or less inevitable. So, for instance, imagine three agents, S1, S2 and S3, each of whom is choosing among acts A, B and C. S1, let's imagine, will suffer a \$100 loss regardless of which of the three he chooses. For him, the \$100 loss is entirely subject to the canceling effect; in deciding what to do, S1 is not under pressure to grant the \$100 loss any reason-giving force. S2, by contrast, will suffer the \$100 loss only if he performs act A or B, but not if he performs act C; S3 will suffer the loss only if he performs A, but not if he performs B or C. Since part of what is involved in deliberating among three different actions is pair-wise comparison — all our agents must decide if they prefer A to B, A to C and B to C — the \$100 loss is somewhat subject to the canceling effect in S2's deliberations, and not at all in S3's. When S2 is determining whether to rank A above B or vice versa he is under no pressure to grant the \$100 loss reason-giving weight; since it attaches to both actions, it is not relevant to the question of which of the two to prefer. He is, however, under pressure to grant the loss weight in comparing A to C and in comparing B to C. Thus, for him, the loss is somewhat subject to the canceling effect. It is subject to it to a greater degree than it is for S3 and to a lesser degree than S1.

Given that the canceling effect is a matter of degree in these two different ways, what we should really say is that a punishment is deserved only if the agent is under pressure to grant the illegality of the act *some sufficiently high degree* of reason-giving force in some prior deliberations. If the act's illegality is strongly enough subject to the canceling effect that the punishment can no longer be said to be brought on the agent by himself, then the agent is not deserving of punishment. I am not in a position to say what degree of canceling, in either sense, is enough to undo desert. I'm not sure, in fact, that anything satisfactory can be said on that score. However, for our purposes the important point will only be that there is some degree of reason-giving force that the agent is under pressure to assign the act's illegality in some prior deliberation, if the punishment for that act is to be deserved.

To solve the problem of private entrapment it is sufficient to show the following: (1) The illegality of the act of an unpredisposed agent who is tempted by the government is sufficiently subject to the canceling effect to undermine desert of punishment for that act; (2) The illegality of the act of an unpredisposed agent who is tempted, not by the government, but by a

private party, is not sufficiently subject to the canceling effect to undermine desert of punishment; and (3) A predisposed defendant was subject to sufficiently strong pressure to grant reason-giving force in his deliberations to the act's illegality to satisfy the necessary condition for desert specified above, regardless of whether he is tempted by the government or a private party. (1) provides an explanation for why unpredisposed defendants tempted by the government are not deserving of punishment. (2) provides an explanation for why the same excuse does not reach to those tempted by private parties. And (3) explains why predisposition defeats the entrapment defense.

Since (3) is the easiest of these three claims to establish, I begin with it. If a defendant is predisposed in the sense outlined in the Subjective Test, then in order to determine if he is deserving of punishment we need to look at more than just the deliberations that led to his decision to give in to the government's temptation. After all, if there is evidence of predisposition, then there is evidence of prior deliberation about whether or not to perform an act of the sort that he later did perform in response to the government's temptation. And, importantly, those prior deliberations are not disconnected from the final deliberations on which he acted; rather, the deliberations launched in response to the government-supplied temptation are a continuation of deliberations begun earlier. (This is one of the results reached in section 2.) But there is reason to think that he was subject to pressure to grant the act's illegality reason-giving force in those earlier deliberations. Given that those earlier deliberations are part and parcel of the same deliberative process on which he acted, it is right to say that he satisfies the necessary condition for desert specified above. If we punish him, we might very well be in a position afterward to say, in justification for our act, "He brought it on himself." Thus, the relevance of predisposition is simply that a predisposed defendant was considering acts of the sort he committed, was under pressure to grant the illegality of those acts reason-giving force in his deliberations, and, as part of the same deliberative process, chose those acts despite such pressures. This still leaves open the possibility, to be closed shortly, that an unpredisposed defendant, tempted by the government, is just as deserving of punishment.

Before moving on to claims (1) and (2), let's return, briefly, to the case of *Gendron*. Gendron, recall, like Jacobson, bought child pornography through the mail and was subject to essentially all of the same treatment from the government as Jacobson. However, as he was seeking child pornography from the first and appeared to have an entirely prurient interest in obtaining it, there is sufficient reason to think he was predisposed. We're now in position to see why his entrapment defense is to be denied. It is to be denied not because of the prurience of his motives for purchasing child pornography — and this a good thing, for had Jacobson been buying porn from a real company that really did fund free-speech lobbying with the proceeds, he would

have been entirely guilty of a crime despite the relatively benign nature of such motives. No, Gendron is to be punished because there is good evidence to suggest that he brought the punishment on himself. There is good evidence, that is, that well before the government entered the scene, Gendron had been considering the purchase of child porn and had been under legal pressure to grant the illegality of that action reason-giving weight in his deliberations. Prurient interests in child porn do not spring up overnight and, even if they did, they would not inspire a man to seek ways to buy in bulk. It also seems clear from the way in which Gendron responded to the government's efforts to tempt him, that the deliberation in which he engaged was a continuation of those prior deliberations, and he elected to buy banned pornography despite the pressure to grant the act's illegality reason-giving weight. We have identified the relevance, then, of Gendron's predisposition: It tells us that he deliberated in the way necessary to be justly subject to punishment.

Now, consider claim (1) — the claim that the illegality of the act of an unpredisposed agent tempted by the government is sufficiently subject to the canceling effect to undermine desert. If a defendant is not predisposed, then for our purposes the only deliberation relevant to the question of whether or not he was deserving of punishment is the deliberation that issued in the decision to give in to the temptation supplied by the government. The reason is this: If we accept that he was not predisposed, then we think that one of two things is true. Either we think that he would not, in the ordinary course of things, ever have *considered* performing acts of the sort which he performed in response to temptation, or we think that, if he did consider such acts, and found himself under pressure to grant the illegality of those acts reason-giving force in his deliberations, he would not have chosen the punishment, in the weak sense: He would have given the illegality of the act reason-giving force in his deliberations but he would have elected an action that was not illegal. The former sort of unpredisposed defendant never would have faced the pressures to grant the act's illegality reason-giving force in his deliberations, and so would never have responded to those pressures in the way required to be deserving of punishment. The latter sort of unpredisposed defendant would have faced such pressures, but would have responded as we think a citizen ought, and so would not have acted so as to warrant punishment. Hence the question is only whether, *in deciding to give in to the government's temptation*, he brought the punishment on himself. Was he subject to pressure to grant the illegality of the act reason-giving force in those deliberations? No deliberations prior to those are of relevance to the question that the court must answer.

Now notice that the government is not supplying the defendant with the temptation idly; rather, it is doing so *for the purpose of causing him to perform an illegal act* should he give in to the temptation. What this means, and it is

nicely illustrated in the most important entrapment cases, is that the government tracks performance by the defendant of an illegal act. By “track” here I mean that the government tracks this result in the same sense in which a heat-seeking missile tracks a heat source: Should the heat source move, the missile will move also, and, similarly, should the agent resist accepting a temptation to perform an illegal act the government will adjust the form of the temptation so as to increase the chances that he will. The government in such cases is striving to find the right temptation, the one that will induce the defendant to give in. If it finds that one approach does not work, that the defendant resists, then it will try another avenue. In *United States v. Woo Wai* (233 F. 412 (1915 U.S. App.)), for instance, the case in which INS agents convinced the defendant to smuggle illegal immigrants, the government made repeated efforts to induce Woo Wai to comply and gave him more and more plausible assurances that he would not be caught. Similarly, in *Jacobson*, the government tried a wide variety of different tactics to induce Jacobson to order child pornography through the mail, including, in the end, offering him the pornography as a way of supporting free-speech lobbying efforts. The point is that the government in such cases is ready to tailor the temptation to the defendant in order to find the form of temptation to which the defendant is willing to submit. In these cases, then, a third party, aware of the force that the government was applying to the defendant, might be able to see, for instance, that when Woo Wai is choosing on an early occasion whether or not to accept the temptation, down the road of acceptance lies immediate violation of the law, while down the road of rejection lies a more attractive offer and violation of the law in giving in to it. If the government were all-powerful and completely determined to cause the defendant to give in to a temptation to perform an illegal act, and if the defendant’s entrapment defense was to be rejected, then every act that the defendant was in position to consider would be illegal. Thus, under these, admittedly stringent conditions, the act’s illegality is completely subject to the canceling effect. Given that it is the defendant’s illegal conduct at which the government aims, and given the imagined unlimited capacity to provide temptation, eventually the government will get what it wants if the defendant’s entrapment defense is not honored. Given more reasonable assumptions — the government has limited resources and only so much interest in the defendant — the fact that the government tracks the defendant’s performance of an illegal act makes the act’s illegality more subject to the canceling effect in the defendant’s deliberations (in the second sense in which subjection to the canceling effect is a matter of degree — it is more inevitable) than it would have been in the absence of that tracking.³²

The point just made — that the government tracks the performance of an illegal act by the defendant — lies at the center of the solution to the problem of private entrapment, and so it is worth explaining in a bit more

detail.³³ To understand the point, consider Figure 1, which represents, in two different ways, the decision that the defendant, unbeknownst to him, faces. Consider, first, the extended form representation of that decision. At first, the defendant must decide whether or not to accept T1, the first temptation issued to him by the government. If he accepts it, he performs an illegal act. But if he rejects it, under the assumption that the government is tracking his performance of an illegal act, he will face a new temptation to perform an illegal act; if he accepts that temptation, he performs an illegal act; if he rejects it then, again, he faces a new temptation, and so on. Assuming the government is truly all-powerful and truly determined to bring it about that the defendant acts illegally, the defendant can never hope to gain the goods promised by the act in question, the goods that make the act tempting to him, through the avenue he believes available to him, without thereby performing an illegal act. This is why there are no branches that involve both the acceptance of a temptation and his performance of a legal act.

Notice, however, that it is not entirely clear from the extended form representation of his situation why the illegality of the act in question is subject to the canceling effect in his deliberations. After all, when deciding whether or not to accept T1, the agent is choosing between performing an illegal act and refraining from such a performance; he does not seem to be faced with a choice between illegal acts. However, the normal form representation of the defendant's situation makes it clear that this description of the choice he faces is misleading. In the normal form representation, the defendant is represented as choosing among temporally extended courses of conduct, or paths from the starting point to a terminal node. He is represented as choosing among (1) accepting T1, (2) rejecting T1 and accepting T2, (3) rejecting T1 and T2, but accepting T3, etc. However, again under the assumption that the government is tracking his performance of an illegal act, every course of conduct that involves his receipt of the goods promised by performance culminates with his performance of an illegal act. So, if we construe the agent as choosing, from the first, between extended courses of conduct, it is clear why the illegality of the act is subject to the canceling effect: Every temporally extended course of conduct that he is in position to consider culminates in his performance of an illegal act.

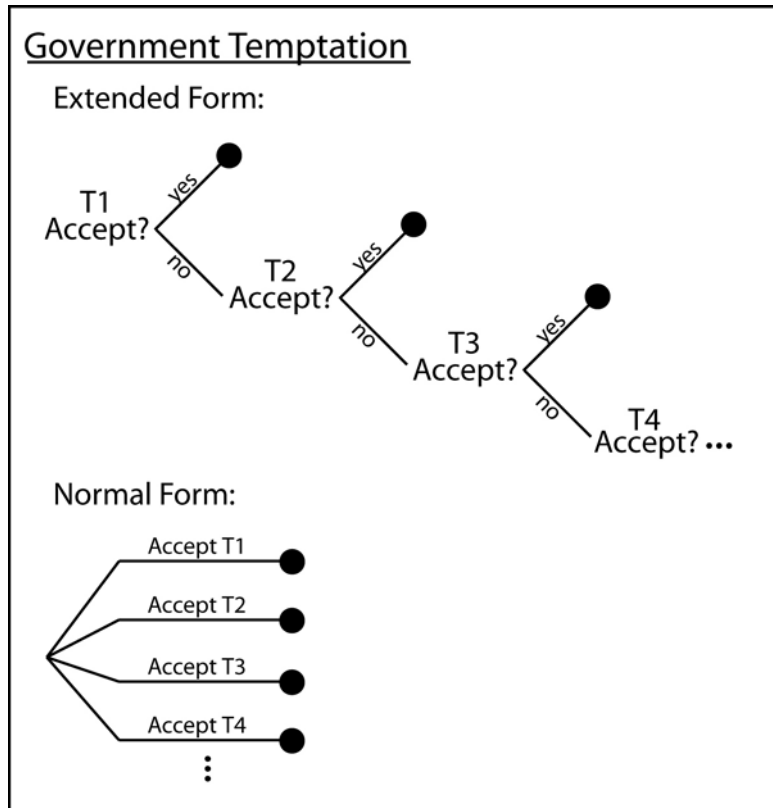


Figure 1: T1, T2, etc. represent various temptations the government is prepared to issue to the defendant. Black circles represent the performance by the defendant of an illegal act.

This raises a question: Why, in thinking about the pressures on deliberation faced by the agent, should we consider the normal form representation of his decision rather than the extended form representation? A first point to make is that, even under the extended form representation, it can appear that the illegality of the act is subject to the canceling effect. After all, we might construe the agent as under pressure, when considering what will happen if he rejects T1, to take into consideration what his future selves will do given the choices they will face. It would be irrational for him, we might say, to imagine that there is no more to the question of what to do than what to do now; it would be irrational to be entirely temporally myopic. Or, to put the point a slightly different way, we might say that the relevant question before him when facing T1 is what the conditional probability is that he will act illegally if he accepts the temptation — roughly, 100 percent — in comparison to the conditional probability that he will act illegally if he rejects the temptation — which, given that the government is tracking his performance and is determined to find the right temptation, is, also, roughly 100 percent.³⁴ Since the two conditional probabilities are the same, the illegality of the act is subject to the canceling effect. These are reasons for thinking that no rationally relevant information is represented in the extended form diagram that is not also represented in the normal form diagram; the fact that each choice the agent makes is between performing an illegal act and refraining from performance doesn't capture the nature of the pressures on his delibera-

tion. However, even if there is a mistake here — even if extended form is not reducible to normal form — there is reason to think that the relevant pressures on deliberation for our purposes here are those that are apparent from the normal form representation. The reason is that an agent can surely not be criticized for choosing between temporally extended courses of conduct. Perhaps he need not; perhaps he would comply with all of the pressures on deliberation by construing his choice in some other way that is more adequately represented in extended form. But for our purposes the question is whether or not the agent is under pressure to grant the illegality of the act reason-giving weight; if there is a way for him to respond appropriately to all of the relevant pressures on deliberation without doing so, then his deliberations have not failed in the way required for desert of punishment.

If the foregoing is right, a judge trying to decide whether or not to grant a defendant's entrapment defense in a case in which it is clear that the defendant is not predisposed ought to reason like this: If the defense is denied, then the act is illegal; thus, to deny the defense is to say that punishment is dictated by the law in this circumstance. But if so, then, given the government's willingness to do what was necessary to get the defendant to give in, a wide range of actions considered by the defendant were illegal. But then the illegality of the act was subject to the canceling effect over that set of actions: The defendant was not under as strong legal pressure to grant it reason-giving weight as he would have been had the government not tracked his performance of an illegal act, since granting the act's illegality full weight would not alter the comparative degree of support by reasons given to the various acts he was considering. But then the punishment was not chosen in the weak sense, and so would not be brought on the defendant by himself, were it to be issued. Therefore, punishing him is appropriate only if he is undeserving of the punishment. It follows, given the retributive justification of punishment, that he is not to be punished for the act and his entrapment defense is to be honored. Unpredisposed defendants, induced to act through a temptation issued by the government, are not deserving of punishment.

It may be clear already why this reasoning does not equally apply to a defendant who is tempted by a private party, rather than the government. That is, it may already be clear what the argument is for (2) — the claim that the illegality of the act of a defendant tempted by a private party, rather than the government, is not sufficiently subject to the canceling effect to undermine desert of punishment. Private parties do not ordinarily track the performance by an agent of an illegal act, even when they track his compliance with the temptation. Ordinarily, that is, a private party may be willing to make more and more attractive and tempting offers to commit a crime, but he does not do so just in so far as such offers make more likely the possibility of the victim's thereby doing something illegal. There may be coincidence between a private party's sweetening of the pot and the likelihood of the victim doing something illegal, but there is no *design* in that connection in the

way that there is when the government is doing the tempting. What the absence of design implies is that when a private party is doing the tempting there will be possible temptations that he is willing to issue, the acceptance of which will not involve illegal conduct. To see the relevance of this point, consider the set of illegal actions that are considered by the agent who is tempted by the government and compare it to the analogous set of actions considered by the agent who is tempted by a private party. The latter set is usually smaller than the former for were a way of legally giving in to the temptation to appear, the government would remove the temptation, rescind the offer for that act, where the private party would not. The private party doesn't care whether or not the act is illegal, while the government does. In fact, were the private party to recognize that an obstacle to the agent's acceptance of the temptation is the illegality of the act he would thereby perform, the private might endeavor to find a way for the agent to accept the temptation without acting illegally. What this means is that the private party's aims do not make the illegality of the defendant's act nearly as strongly subject to the canceling effect as the government's aims do.

To see the point, consider Figure 2. Assume that the private party, like the government in our earlier example, is determined to get the defendant to accept a temptation. But assume, also, that the private party tracks not the performance of an illegal act on the part of the defendant, *but the performance of some act that happens to be illegal*. Imagine, for instance, that a company pedaling child pornography received Jacobson's name from the same list that the FBI used. They, just like the government, send him catalogs in order to tempt him to commit a crime, a crime that will put money in their pockets. Now imagine that, after a year or so of regular mailings, the law against child pornography is changed so that only half, instead of all, of the materials advertised in both the government's and the private party's catalogs are illegal; perhaps the age at which a person is to be considered a minor is lowered. The government immediately alters the catalogs that they send to Jacobson so that only illegal materials are advertised. Given that it is in the business of causing Jacobson to do something illegal, what would be the point of continuing to advertise materials that Jacobson has a legal right to order? What the government is trying to do by making this change is to ensure that as few as possible of the options available to Jacobson are legal. By contrast, the private party continues to send Jacobson the same catalogs as before; the private party can make the same dollar regardless of the legality or illegality of Jacobson's selections. In Figure 2, imagine that T_n is the temptation that the private party issues after the laws against child pornography are changed; he is providing Jacobson an opportunity to do as he has been tempting him to do all along, but to do so legally. The illegality of Jacobson's act, then, is subject to the canceling effect to a greater degree when he is tempted by the government than when he is tempted by a private party; in

the former case there are a greater range of actions with respect to which his deliberations would be unaffected were he to grant no reason-giving force to the act's illegality. Importantly, this result is not specific to this particular example. So long as the government is, and the private party is not, *aiming* at the performance by the defendant of an illegal act, there will be possible temptations that the government would not issue and the private party would and that could be accepted without doing anything illegal. This is the solution to the problem of private entrapment: Private parties, since they do not track the illegality of the acts of those whom they tempt, do not make it the case, as the government does, that the illegality of the act is subject to the canceling effect to the degree necessary to make the defendant undeserving of punishment.

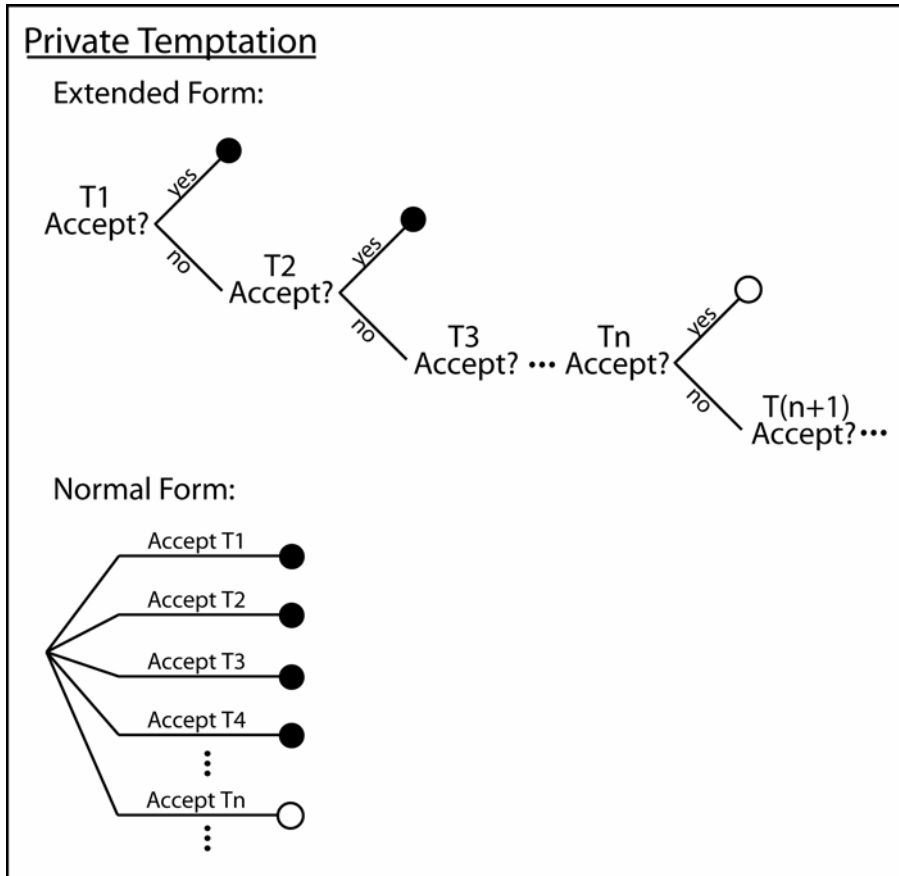


Figure 2: T1, T2, etc. represent various temptations the private party is prepared to issue to the defendant. Black circles represent the performance by the defendant of an illegal act, while white circles represent performance of a legal act.

Notice that, as described, the difference between the government and the private temptor is a difference in what they would actually do at some point in the future of the temptation that the defendant actually accepted.

Now this difference is entirely *counterfactual*; the defendant actually accepts T1, the very first temptation issued to him. What the diagrams in Figures 1 and 2 represent are *possibilities*. We look at the situation that the defendant faced when he accepted the temptation and we ask what he could have expected under various conditions. What would have happened if he'd rejected the first temptation? What would have happened if circumstances downstream of that decision were in various ways different from what we expect them to have been? What we find is that, when the government is involved, more of the answers to these hypothetical questions involve description of worlds in which the defendant acts illegally than when a private party is doing the tempting. The outcome of acting illegally is more counterfactually robust — it occurs across a greater range of relevant possible worlds — when the party doing the tempting tracks the performance by the agent of an illegal act, as the government ordinarily does. What follows is that the illegality of the act is more subject to the canceling effect when the government tempts than when a private party does.

I have not provided an argument for thinking that this difference in degree of subjection of illegality to the canceling effect — the government's actions make it more subject, the private party's less so — is enough to make a difference in desert. Why should we think that the defendant tempted by the government is under *sufficiently* less pressure in deliberation to warrant the claim that he did not, and the defendant tempted by a private party did, bring the punishment on himself? I don't have an answer to this question any more than I have an account of what degree of pressure to grant reason-giving force to the act's illegality is necessary to warrant the thought that the defendant brought the punishment on himself. Still, the theory on offer bears sufficient fruit even without an answer to this question. What it provides is an explanation for why there should be a difference in culpability between those tempted by the government and those tempted by private parties, even if it doesn't show that always and in general this difference will be found. There are undoubtedly cases in which the government is not terribly zealous — they don't track the defendant's performance of an illegal act very well; and, conversely, there may be cases in which private parties really do track such performance. If there are such cases, then they would fall on the opposite side of the line than one might initially expect.

It's important to highlight an admission that has just been made. On the view presented here, it is the tracking of illegal action on the part of a temptor that undermines the tempted's culpability for compliance. It is a *contingent matter* that the government tracks in this way and private parties do not. To some, the contingency of connection between government temptation and the kinds of temptation that undermine desert of legal punishment appears to undermine the solution to the problem of private entrapment on offer. If the position advocated here is right, shouldn't we have a "tracking"

defense, rather than an entrapment defense? Shouldn't defendants have to show that whoever supplied the temptation they accepted was tracking their performance of an illegal act?³⁵ The question to ask about this objection, however, is not whether it traces out the implications of the argument here correctly — it does — but whether the results the objector points to amount to *objections*. I explain.

The solution to the problem of private entrapment offered here does imply that a party who tracks performance of an illegal act by the defendant with sufficient zeal does make the illegality of the act subject to the canceling effect and so undermines the defendant's culpability for it. Since culpability is necessary for legal responsibility, it follows that some who accept temptations issued by private parties ought not to be convicted. Imagine, for instance, a child advocacy organization that, fed up with the FBI's paltry efforts, takes it on itself to tempt would-be consumers of child porn into purchasing such materials so as to make prosecution of them possible. On the view defended here, if such an organization really tracks the performance of an illegal act on the part of the defendant, and if it really does it with sufficient zeal to strongly subject the act's illegality to the canceling effect, the defendant is not culpable for having accepted the temptation. What this implies is that the fact that it is the government that issued the temptation is not of *substantive* importance to the defense, but of profound *probative* importance. That is, given the way the police does business, and given how much power they have, we think it extraordinarily likely that a government temptor tracks performance of an illegal act on the part of the defendant. And given how common it is to find private parties who want others to do things that happen to be illegal, but don't particularly care *that* they are illegal (think of drug traffickers, for instance), it is extraordinarily likely that private temptors do not track illegal action *per se*. Thus, the fact that the government issued the temptation, rather than a private party, is of overwhelming evidentiary significance. It is significant enough, in fact, as to make the showing of that fact alone sufficient under certain circumstances to grant the defense. When an unpredisposed defendant shows that the government tempted him, he shows it to be so likely that his culpability was undermined as to make the risk of wrongful conviction far too high for the government to proceed with conviction and punishment. No comparable risk is present when a private party tempts, and so there is no defense of private entrapment.

Return, briefly, to the disagreement between advocates of the Subjective and Objective tests. Nothing said here can adjudicate the dispute between these two parties. Rather, both can benefit from the solution to the problem of private entrapment just offered. It is obvious, perhaps, how an advocate of the Subjective Test can use the argument just offered: If the argument works, it shows that the defense is to be granted just when the Subjective Test says it should be, for it shows that just those defendants who are

tempted by the government and unpredisposed are not culpable for their acts. But, in fact, the argument can be accepted also by someone who defends the Objective Test, particularly in the form offered in the *Model Penal Code*, according to which, recall, government temptation amounts to misconduct just in case it risks ensnaring the unpredisposed. Someone who advocates this position can agree that what's bad about risking ensnaring the unpredisposed is that when the government succeeds in ensnaring such a person, it ensnares someone who's not culpable. And such a person can agree that someone who is ensnared by a private party, even if unpredisposed, is still culpable for the reasons offered here. And, further, such a person can agree that when the government ensnares the predisposed it ensnares someone culpable. But the advocate of such a position can further insist that none of this means that we ought to grant the defense only to the unpredisposed. In fact, it might be argued, it is so bad to ensnare the unculpable, that we're going to grant the defense anytime the government even risks that, and even in cases in which, we admit, the defendant is culpable. The argument between the advocate of the Subjective Test and someone who accepts the *Model Penal Code's* test, supplemented by the offered solution to the problem of private entrapment, then, reduces to a debate very close in structure to the debate over the exclusionary rule: Both sides admit that, for instance, the practice of excluding coerced confessions from evidence results in some guilty people going free, but one side, and not the other, thinks this is worth it to minimize the risk of false conviction on the strength of coerced confessions. Thus, the debate between advocates of the Subjective and Objective tests can't be settled by appeal to the views' respective abilities to solve the problem of private entrapment, but must be settled, if it can be, through consideration of the normative question of what limitations on government power are worthwhile despite the fact that they result in less effective law enforcement.

Conclusion

It is worth highlighting two results, both of which are strongly suggested by the central claim of this paper, the claim that there is a difference in culpability between those who give in to government-supplied temptation, and those who give in to temptation supplied by a private party. The first result is this: Repeatedly, courts identify one particular feature of government conduct as objectionable in explaining their decision to honor an entrapment defense: the government's dogged persistence to cause the defendant to act in the way banned by the law.³⁶ But why should it matter that the government is persistent in its efforts? Private parties, too, can be persistent in their efforts, yet we don't grant a defense to those who act criminally as a result of such persistence. The discussion here explains why governmental persistence,

in particular, matters: The government is not simply persistently trying to get the entrapped to act in a certain way; it is trying to get him to *act criminally*. What it is persistent *about* is the criminality of the defendant's conduct. But what this means is that the illegality of the defendant's act becomes inevitable for that defendant in a way that undermines the defendant's deserving of punishment.

The second, and more philosophically important moral to be drawn from this discussion, concerns the role of mental states in culpability assessments. It is natural to think that a defendant's legal culpability is entirely a matter of the mental states of the defendant, the mental states that in some way contributed to the performance of the acts of which the defendant is accused. (Such an assumption is apparent in the very naming of the Subjective Test, as though in assessing culpability we need only consider those facts that are immediately accessible to the subject, namely mental states, which are not so easily accessible from a third-person point of view.) But if the conclusion of this paper is correct, this is a mistake. Governmentally and privately tempted defendants have exactly the same sorts of mental states; both have their reasons for wanting to do something that the law bans, and both, typically, believe themselves to be given an opportunity to act criminally by a private party. Since, as has been argued, they differ in culpability, culpability is not solely a matter of mental state.³⁷ This result, in turn, has implications for the ongoing philosophical project of specifying necessary and sufficient conditions for the possession of the sort of freedom needed for responsibility. If there is more to culpability than mental state, then it seems unlikely that an analysis of the concept of freedom can be adequate if it appeals only to the mental states of the free agent.³⁸ Theorists concerned with the so-called "free will problem," then, would do well to turn their thoughts to the foundations of the entrapment defense and the inevitable problem of private entrapment.³⁹

¹ *Genesis* 3:13

² For an extremely helpful overview of the historical development of the entrapment defense, see Rebecca Roiphe, "The Serpent Beguiled Me: A History of the Entrapment Defense," 33 *Seton Hall L. Rev.* 257 (2003).

³ There is no "problem" of private entrapment if privately entrapped people ought to be excused under just the same conditions as the governmentally entrapped. This is claimed in Ronald J. Allen, Melissa Luttrell and Anne Kreeger, "Clarifying Entrapment," 89 *J. Crim. L. & Criminology* 407 (1999). Allen, Luttrell and Kreeger claim that the entrapment defense should be granted just in case the temptor has offered higher than market-level inducements to commit the crime. Since either a private party or the government can provide such inducements, a defendant's responsibility is undermined in either case. It seems to me, however, that the right response to this is to affirm *modus tolens*: Given that the privately entrapped ought not to be granted a defense, it follows that defendants shouldn't be excused simply on the grounds that they received higher than market-level inducements to perform.

⁴ Concern with the problem can be seen in recent cases as well. For instance, in *United States v. Gendron* (18 F.3d 955 (1994 U.S. App.)) the point is raised bluntly: "[W]ere the issue simply the defendant's state of mind...the law would permit an innocent minded defendant to raise an entrapment claim when a private person 'induced' him...to commit a crime. But the law does not authorize the defense in those instances, however 'outrageous' the private person's conduct." See also *United States v. Russell* (411 U. S. 423 (1973)), and *Sherman v. United States* (356 U. S. 369 (1958)). In both of those cases the minority objects to the majority's ruling by appeal to the problem of private entrapment. Relatedly, Christopher D. Moore argues that the problem of private entrapment constitutes a profound difficulty for the entrapment defense. See, Christopher D. Moore, "The Elusive Foundation of the Entrapment Defense," 89 *Nw. U.L. Rev.* 1151 (1995), esp. p. 1157.

⁵ This is not the first paper to try to solve the problem. Roger Park tries to solve the problem by arguing that the privately entrapped and the governmentally entrapped are equally innocent, but that there are reasons of public policy for denying the defense to the privately entrapped. (See Roger Park, "The Entrapment Controversy," in *Minnesota Law Review*, v. 60, 1976.) Park's position on this issue is convincingly attacked in Andrew Altman and Stephen Lee, "Legal Entrapment," in *Philosophy and Public Affairs*, v. 12 (1), 1982, pp. 51-69, esp. pp. 58-9. For further discussion, see Michael Gorr, "Entrapment, Due Process and the Perils of 'Pro-Active' Law Enforcement," in *Public Affairs Quarterly*, v. 13 (1), 1999, pp. 1-24, esp. p. 13. The primary problem with Park's approach, as Altman and Lee point out, is that the fact that a person is privately entrapped does not serve to morally exculpate quite independently of the particular policy-entwined purposes of a legal defense. In other words, there appears to be something peculiarly appropriate about convicting the privately entrapped, and not the governmentally entrapped, even when we confine our gaze to particular cases and not, as we do when bringing in policy considerations, to the aggregate of cases. A true solution to the problem of private entrapment needs to explain why the privately entrapped are not, and the governmentally

entrapped are, rightly subject to legal punishment quite independently of extrinsic policy considerations.

⁶ On occasion, the line between government agent and private party is not so clear. For instance, in *United States v. Manzella* (791 F.2d 1263 (1986 U.S. App.)), a DEA agent convinced a man named Rizzo, who had no connection to the government, to sell him some cocaine. Rizzo brought the deal to the defendant, Manzella, who sold the cocaine to the agent, with Rizzo acting as intermediary. Manzella claimed that he was entrapped. The court rejected the defense on the grounds that Rizzo was not working for the government.

⁷ Normally, in order to show that he has been entrapped, a defendant must show that he was "induced" to act by the government. Sometimes the term "inducement" is used in an evaluatively loaded way to mean not just that the government tempted the defendant into acting, but that in doing so it acted badly or inappropriately. Since, as we will see, one of the central questions about entrapment is whether or not it is important to know if the government acted badly toward the defendant, it is important to have an evaluatively neutral term to capture the role the government plays in all the cases that concern us. For our purposes, then, to say that the government tempted the defendant to perform the crime, and that the defendant acted as he did in response to its temptation, is not to imply anything about whether or not the government overstepped its bounds in issuing the temptation.

⁸ As Kyron Huigens helped me to see, the term "culpable" is sometimes used differently from the way it is being used here. It is sometimes used as a status term, to refer to those who are, we might say, normatively competent or who are possessed with the basic set of capacities that qualify one to be asked to meet legal demands. As the term is being used here, however, it picks out agents only in so far as they stand in a certain (somewhat obscure) relation to a particular act. The term is being used here in the sense of "culpability for" a particular act and not as a way of identifying a particular status of the agent independently of any particular action.

⁹ These cases are: *Casey v. United States* (276 U. S. 413 (1928)), *Sorrells v. United States* (287 U. S. 435 (1932)), *Sherman v. United States* (356 U. S. 369 (1958)), *United States v. Russell* (411 U. S. 423 (1973)), *Hampton v. United States* (425 U. S. 484 (1976)), *Jacobson v. United States* (503 U.S. 540 (1992)).

¹⁰ The distinction appears starkly in *Sorrells v. United States* (287 U. S. 435 (1932)). Writing for the majority, Justice Hughes says that the "controlling question" is "whether the defendant is a person otherwise innocent whom the Government is seeking to punish for an alleged offense which is the product of the creative activity of its own officials." And thus, he concludes, the defendant who adopts an entrapment defense makes himself subject to a "searching inquiry into his own conduct and predisposition." By contrast, Justice Roberts, writing for the minority, says that cases of entrapment involve a "prostitution of the criminal law" on the part of the government and that to grant an entrapment defense is in fact to judge a defendant to be "not guilty by reason of someone else's improper conduct."

¹¹ In fact, talk of "predisposition" did not enter until after *Woo Wai* was decided. Rather, in *Woo Wai*, the court ruled that given that Woo Wai would not have, in the ordinary course of things, encountered circumstances that would have prompted him to commit similar acts, that he was "otherwise innocent" prior to being subject to temptation by the government. For our purposes, however, the point is the same.

¹² In *United States v. Gendron* (18 F.3d 955), the court proposes a general test for determining if a possible circumstance is "relevant." Writing for the majority, Judge Breyer says: "The right way to ask the question [viz. Subtracting consideration of the government's conduct, would the defendant have otherwise committed the crime?], it seems to us, is to abstract away from — to assume away — the present circumstances insofar as they reveal government overreaching. That is to say, we should ask how the defendant likely would have reacted to an ordinary opportunity to commit the crime." Breyer's idea is that a counterfactual occasion is relevant so long as it does not involve government misconduct. If this test is followed to the letter, then every entrapment defense will be denied. After all, we know that the defendant would accept the temptation the government supplied even if it hadn't been supplied by the government; the defendant didn't know that it was the government who was tempting him. And, crucially, an occasion in which he is tempted just as he was by a private party is a "relevant" occasion by this standard since it does not involve any government misconduct. It thus follows that the defendant is predisposed by this standard just by virtue of the fact that he accepted the government's temptation.

¹³ Thus, what will be argued here is in disagreement with the position of the *Model Penal Code*, which claims, "[T]he defendant whose crime results from an entrapment is neither less reprehensible or dangerous, nor more reformable or deterrable, than other defendants who are properly convicted." (*Model Penal Code*, Commentary on §2.13, v. 1, p. 406). I agree with the *Model Penal Code* that the defense cannot be justified through appeal to the lack of "dangerousness" (special deterrence), "deterrability" (general deterrence), or "reformability" (rehabilitation), of those whom it excuses. I disagree, however, that the defense cannot be justified through appeal to the relative lack of "reprehensibility" (retribution) of those whom it excuses. That, in any event, is what is to be argued here.

¹⁴ Thanks to Scott Altman, Greg Keating and Andrei Marmor for, in different ways, pushing lines of thought of this sort.

¹⁵ Jonathan C. Carlson also tries to justify the entrapment defense on retributive grounds. (See Jonathan C. Carlson, "The Act Requirement and the Foundations of the Entrapment Defense", 73 Va. L. Rev. 1011 (1987).) Carlson claims that in cases in which the entrapment defense is to be honored, no one is harmed, and so no wrong deserving of punishment has been committed. If this justification for the defense were adequate, however, it would justify excusing any defendant who is tempted by the government, including the predisposed, assuming that the government's involvement prevented any harm from being caused. In addition, if such a justification works for entrapment, it undermines the justice of any punishment in cases of "victimless" crime. These are unsatisfactory results.

¹⁶ This point is raised by the minority in *Sorrells v. United States* (287 U. S. 435 (1932)): "The [Subjective Test], in effect, pivots conviction in such cases, not on the commission of the crime charged, but on the prior reputation or some former acts of the defendant not

mentioned in the indictment." See, also, discussion in B. Grant Stitt and Gene G. James, "Entrapment: An Ethical Analysis," in *Moral Issues in Police Work*, F. Elliston and M. Feldberg eds., Allentown: Rowman and Allanheld, 1985, pp. 129-145, esp. p. 133. Feinberg discusses a closely related concern. See Joel Feinberg, "Causing Voluntary Actions," in *Doing and Deserving*, Princeton University Press: Princeton, 1970, pp. 175-6.

¹⁷ Joel Feinberg makes the following remark: "[T]he predisposed criminal is no less guilty for being induced to act by a policeman in a disguise than thousands of others of like disposition who are induced by private persons." (Joel Feinberg, "Causing Voluntary Actions," in *Doing and Deserving*, Princeton University Press: Princeton, 1970, p. 176) Feinberg is not quite giving voice to the claim that the Subjective Test faces the problem of private entrapment, for he is insisting on the fact that there is no difference between those the Subjective Test condemns (the predisposed) and those who are tempted by private parties and who our law denies the defense of entrapment. By contrast, the problem of private entrapment arises because it appears that there is no difference between those the Subjective Test excuses (the unpredisposed) and those who are privately entrapped. It is this latter appearance, and not the one to which Feinberg points, that, if more than an appearance, threatens the Subjective Test. George Fletcher stresses the point: "There is no single theory of the [entrapment] defense that can explain both (1) the exclusion of entrapment by private parties and (2) the limitation to suspects not 'predisposed' to commit the offense." (George Fletcher, *Rethinking Criminal Law*, Oxford University Press: Oxford, 2000, p. 543.) Section 4 of this paper offers just the sort of theory that Fletcher denies to be possible.

¹⁸ In fact, the *Model Penal Code* justifies its choice of an Objective Test through appeal to the problem of private entrapment: "Defendants who are aided, solicited, deceived or persuaded by police officials stand in the same moral position as those who are aided, solicited, deceived or persuaded by other persons; yet no one suggests a general defense for the latter." (*Model Penal Code*, Commentary on §2.13, v. 1, p. 406). In this, the *Model Penal Code* echoes the view expressed by Justice Stewart in his dissenting opinion in *United States v. Russell* (411 U. S. 423 (1973)), in which a plea is made for some version of the Objective Test: "That [the defendant] was induced, provoked or tempted to do so by government agents does not make him any more innocent or any less predisposed than he would be if he had been induced, provoked or tempted by a private person — which, of course, would not entitle him to cry 'entrapment.'" For some discussion, see Michael Gorr, "Entrapment, Due Process and the Perils of 'Pro-Active' Law Enforcement," in *Public Affairs Quarterly*, v. 13 (1), 1999, pp. 1-24, esp. pp. 12-13. Gorr describes the problem of private entrapment as "the single most powerful criticism of any view which includes a subjective element focusing on the defendant's predisposition." (p. 13)

¹⁹ Gerald Dworkin offers a similar version of the Objective Test. (See Gerald Dworkin, "The Serpent Beguiled Me and I Did Eat: Entrapment and the Creation of Crime," in *Law and Philosophy*, v. 4, 1985, pp. 17-39, esp. p. 33.) According to Dworkin, a government-issued temptation is acceptable (and thus the defendant's entrapment defense should be denied) only if the government had probable cause to believe that the defendant was already engaged, or intended to be engaged in criminal activity. Dworkin offers a higher standard for predisposition than that presented in the Subjective Test here — the defen-

dant must actually have been involved in criminal activity, the counterfactual possibility of such engagement is not enough. Also, Dworkin can be thought of as unpacking the *Model Penal Code*'s notion of "risk of ensnaring the unpredisposed"; under Dworkin's view, the government is guilty of creating such a risk just in case it lacked probable cause to believe the defendant to be involved in criminal activity. The differences, however, between Dworkin's approach and that of the *Model Penal Code* are not significant enough, so far as I can see, to provide Dworkin a way of avoiding the criticism of the *Model Penal Code*'s approach offered in the main text.

²⁰ The court creates some confusion by claiming that there is good evidence that Jacobson had a "predisposition to view photographs of preteen sex" (Majority opinion in *Jacobson*). The court, however, is careful to distinguish that predisposition from a predisposition to order child pornography through the mail, the crime of which Jacobson was accused. This confusing aspect of the opinion has generated a great deal of discussion of *Jacobson*. See, for instance, J. Gregory Deis, "Economics, Causation, and the Entrapment Defense," U. Ill. L. Rev. 1207 (2001); Paul Marcus, "Presenting, Back from the (Almost) Dead, the Entrapment Defense," 47 Fla. L. Rev. 205 (1995); Lance Levy, "Criminal Law — The 'Sting' of Government Operations: An Analysis of Predisposition as it Relates to the Entrapment Defense — *Jacobson v. United States*," 26 Suffolk U.L. Rev. 1177 (1992); Brian Thomas Feeney, "Scrutiny for the Serpent: The Court Refines Entrapment Law in *Jacobson v. United States*," 42 Cath. U. L. Rev. 1027 (1993); Nancy Y. T. Hanevitz, "Jacobson v. United States: The Entrapment Defense and Judicial Supervision of the Criminal Justice System," Wis. L. Rev. 1163 (1993); Jack B. Harrison, "The Government as Pornographer: Government Sting Operations and Entrapment," 61 U. Cin. L. Rev. 1067 (1993); Eric Weyand, "Entrapment: From Sorrels to Jacobson — The Development Continues," 20 Ohio N.U.L. Rev. 293 (1993).

²¹ Damon D. Camp claims that there would be something wrong if a hypothetical Jacobson, who was pursuing child pornography aggressively and for prurient reasons, had access to the entrapment defense. (See Damon D. Camp, "Out of the Quagmire After *Jacobson v. United States*: Towards a More Balanced Entrapment Standard," in *the Journal of Criminal Law and Criminology*, v. 83, 1993, pp. 1092-93.) Gendron, real and not hypothetical, is just such a person.

²² For discussion of an earlier draft of the *Model Penal Code*'s test, a discussion which expresses a related criticism to that offered here, see Joel Feinberg, "Causing Voluntary Actions," in *Doing and Deserving*, Princeton University Press: Princeton, 1970, p. 174, n. 22.

²³ This claim is closely related, it seems, to the claim made by Christopher D. Moore to the effect that even the Objective Test, to be plausible, must focus on the defendant's predisposition or lack thereof. See Christopher D. Moore, "The Elusive Foundation of the Entrapment Defense," 89 Nw. U.L. Rev. 1151, esp. pp. 1172-77.

²⁴ My thinking about this was inspired by some remarks of Martin Stone's.

²⁵ The *Model Penal Code* includes entrapment by estoppel by offering the following sufficient (but not necessary) condition for entrapment: A defendant is entrapped if he acted

as a result of the government "making knowingly false representations designed to induce the belief that such conduct is not prohibited." (*Model Penal Code* §2.13(1)(a)).

²⁶ It needn't involve the sort of mistake of law that constitutes a defense under *Model Penal Code* §2.04(3), for instance.

²⁷ This seems to be just what Gerald Dworkin has in mind: "[F]or a law enforcement official to encourage, suggest or invite crime is to, in effect, be saying 'Do this.' It is certainly unfair to the citizen to be invited to do that which the law forbids him to do. But it is more than unfair; it is conceptually incoherent...From the standpoint of one trying to understand and evaluate the system...the conflict is clear." (Gerald Dworkin, "The Serpent Beguiled Me and I Did Eat: Entrapment and the Creation of Crime," in *Law and Philosophy*, v. 4, 1985, p. 32)

²⁸ Thanks to Herbert Morris for the example.

²⁹ Thanks to Mark Greenberg for the example.

³⁰ Why? In broadstrokes, one possible answer is that the different pressures on deliberation derive in the end from different forms of flourishing of our deliberative mechanisms. You ought to consider the means to your intended ends because by doing so you can get closer to realizing a certain ideal: the ideal of the rational deliberator. Similarly, by granting reason-giving force to the immorality or illegality of an act you can get closer to realizing ideals of moral and legal deliberation, respectively. But since there is no reason to expect all of these ideals to be realizable at once, there is no reason to expect reasons of different sorts to cancel one another.

³¹ A consequence can be subject to the canceling effect to some degree, in this first sense, also, when the probability of its occurrence, given performance of a certain action, is less than one. For instance, imagine that act A is guaranteed to result in a \$100 loss, while act B has a probability of doing so of 40 percent. Thus, the expected loss for performing act B is not \$100, but $100 * .4 = \$40$. Therefore, the \$100 loss attaching to act A is somewhat subject to the canceling effect; the agent is under prudential pressure to treat it as a \$60 loss.

³² There is an affinity between the explanation being offered here and the explanation that I give elsewhere for the claim that coercion and indoctrination, as opposed to neutral forces that give the agent just as bad a choice to make, undermine freedom of will. See Gideon Yaffe, "Indoctrination, Coercion, and Freedom of Will" in *Philosophy and Phenomenological Research*, v. 67, n. 2, September 2003, pp. 335-356.

³³ The elaboration to follow owes a great deal to discussion with Scott Shapiro.

³⁴ Thanks to Alex Stein for pointing out the relevance of conditional probability to the point at issue here.

³⁵ Thanks to Andrei Marmor and Martin Stone for independently pushing this objection in conversation.

³⁶ To take just a few examples, see *United States v. Rodriguez* (858 F.2d 809 (1988 U.S. App.)), *United States v. Kessee* (992 F.2d 1001 (9th Cir. 1993)), *United States v. Gendron* (18 F.3d 955 (1994 U.S. App.)).

³⁷ This conclusion is probably consistent with the view of Christopher D. Moore, according to which the entrapment defense is inherently inconsistent with the standard doctrine of excuses. (See Christopher D. Moore, "The Elusive Foundation of the Entrapment Defense," 89 Nw. U.L. Rev. 1151.) If Moore is right that whether or not a defendant has an excuse is a function of his mental state, then the point being made here is the same. For discussion, see Michael Gorr, "Entrapment, Due Process and the Perils of 'Pro-Active' Law Enforcement," in *Public Affairs Quarterly*, v. 13 (1), 1999, pp. 1-24, esp. pp. 13-15.

³⁸ Perhaps the most prominent (not to mention important and insightful) analysis of the concept that makes this mistake is that offered by Harry Frankfurt in a series of connected papers beginning with "Freedom of the Will and the Concept of a Person" in *The Importance of What We Care About*, Cambridge University Press: Cambridge, 1971, pp. 11-25.

³⁹ Portions of this material were presented at the California Institute of Technology, the University of California at Riverside, the UCLA Legal Theory Workshop, the University of Southern California School of Law and Cardozo Law School. Thanks are owed to audiences on all five occasions for helpful remarks. In addition, all of the following people have provided extremely helpful comments: Scott Altman, Michael Bratman, Bill Bracken, David Dolinko, Sharon Dolovich, Gerald Dworkin, Steve Finlay, John Fischer, Peter Graham, Mark Greenberg, Barbara Herman, Pamela Hieronymi, Paul Hoffman, Kyron Huigens, Greg Keating, Andrei Marmor, Herbert Morris, Calvin Normore, Scott Shapiro, Paul Shupack, Alex Stein, Nomi Stolzenberg, Martin Stone and Gary Watson.